ORIGINAL PAPER

Lexicography for loanwords and words with special orthography

Loanwords in Modern Standard Chinese

Christian Schmidt[1] · Chen Jien-shou[2]

Abstract

Loanwords in Modern Standard Chinese have attracted increasing attention in recent years. We try to unravel the complexity that stems from the properties of the Chinese writing system itself, its variant systems (Japanese Kanji), the rich history of the Chinese lexicon, and its interactions with neighboring languages. We argue that the status of a loanword is not binary, but rather dynamic on a developmental continuum, due to the various intentional choices and cumulative efforts involved in loanword integration. We provide a metalanguage for loanword notation to make explicit throughout the study all relevant kinds of relationships between source and target languages during the borrowing process from any language into Chinese. Finally, we show examples for all major loanword categories, including a modified and notated categorization of loanwords in Chinese, based on Shi ([Chinese loanwords] Hanyu wailaici. Commercial Press, Beijing, 2013).

Keywords Modern Standard Chinese · Loanwords · Lexicography · Linguistics

1 Introduction

The Chinese lexicon today counts roughly 20,000 loanwords, many of which have entered the lexicon since the late Qing dynasty. Among these we find roughly 3000 Japanese loanwords (Shi 2019). Huang and Liu (2017) have recorded over 50,000 alphabetic words.

✉ Christian Schmidt
d05142004@ntu.edu.tw

Chen Jien-shou
jienshou@gate.sinica.edu.tw

1   National Taiwan University, Taipei, Taiwan

2   Academia Sinica, Taipei, Taiwan

A good part of this development can be exemplified by how the very definition of what counts as a loanword has changed over the years, which is discussed in Sect. 2.4. We argue that the status of a loanword is not binary, but dynamic on a developmental continuum, due to the various intentional choices and cumulative efforts involved in loanword integration. We introduce a new way of conceptualizing this continuum, i.e. how well loanwords are integrated into the standard vocabulary by measuring their degree of 'Integration' and by how much a loanword has structurally and orthographically changed compared to its source counterpart ('Distance') in Sect. 2.4.2.

We are able to do so due to a notation metalanguage we developed, i.e. a method for making explicit the most important processes during borrowing. We can now elucidate the qualitative differences that are created in the relationship between source and target words.[1]

The article is organized in the following way: try to give a general overview over all major types of loanwords in Chinese in a systematic fashion, that is by focusing on three major aspects of the borrowing process: (a) orthographic integration based on the choice of orthographic representation, (b) phonetic integration, including material borrowing as well as updated and replaced transliterations, (c) semantic transliteration, here focusing on the fundamental differences in terms of morphemic function: iso-morphemic transliteration; various complex transliterations; conceptual remakes, as well as sense extensions. Section 2.4 addresses the aforementioned categories with an admittedly short history of the development of Chinese loanword theories. Here we attend to such questions as: Which categories are mostly contested in the literature? Or, is it reasonable that Japanese loanwords form a separate class of their own? Regarding Japanese loanwords, we take the liberty of shortly discussing a claim by Shen (2017) and others that a considerable amount of so-called Japanese loanwords arguably had been created by Western missionaries and dictionary makers in China, before being transferred to Japan during the Meiji period – evidence pointing to a Western origin of said words. We close this paper with an annotated table of loanword categories, based on (Shi 2013).

2 Methods of borrowing

Methods of borrowing refer to the three basic aspects of a word that can be borrowed: its orthographic properties, sound, and word meaning. These properties can all be selectively borrowed, and selectively combined with autochthonous morphemes. In the outline of this research, we have organized various borrowing strategies under different segments, according to the salience of each aspect's borrowing strategy; however, this simplicity could be misleading. In fact, borrowing is the flip side of the process of word-creation. And word-creation employs the

---

[1] In our case, these differences between source and target word are mostly linguistic, but there are obviously also differences due to the wide societal and historical implications of loanwords. Liu (2008) has argued, for example, that translation is not simply a transliteration between languages, but that we also need to consider the historical context. She argues that language is a historical product. If a word is being extracted from its original context and used in its non-endemic context, its meaning extensions might change drastically, turning the loanword into an entire new word.

entire arsenal of linguistic strategies—to various degrees and successes—to integrate a loanword into the vocabulary. Borrowing is an active linguistic process not only of receiving a loanword, but of actively transforming it, which often involves a high degree of effort on behalf of the 'receiving' language community. This is even more so for Chinese, due to the specific properties of the Chinese script that express form,[2] meaning and sound separately.[3]

   Chinese is different from Western alphabetic languages in that a Chinese word, once communicated via script, will take on one additional layer that gains and actuates some semantic function—the orthographic layer. Due to its orthographic properties a word always has at least one sense associated with its form that may or may not stand independently from the sense that is actuated by its pronunciation, e.g. 'photograph, picture' *zhaopian* is written in standard form as 照$_a$ 片$_1$, but can be humorously written—without changing its pronunciation—as 照$_a$ 骗$_2$, whereas (骗$_2$) means 'to fool s.o.', creating the meaning 'fake photograph'. The whole 2-character combination gains an additional semantic layer due to its potential for exploiting a large number of homonyms in Chinese. In written form, the relationship between a morpheme or a syllable and its representing character is not trivial, but rather a fully self-contained subject matter. The takeaway is that due to its graphic representation, borrowing foreign words into Chinese is more complicated than into alphabetic languages. But it also means that Chinese, by nature of its script, is more sensible to loanwords; hence, more efforts involved in loanword integration.

2.1 Borrowing orthographic properties

The general discussion about loanwords focuses on words. Yet, numerals and scientific units are also some kind of 'words', just written in another character set—the set of numerals or specially assigned letters.[4] Hence, in the following subsections we discuss words with special or mixed orthographies. Section 2.1.1 explores alphabetic words; in 2.1.2 we argue that Arabic numerals also count as (written) loanwords in Chinese, also mentioning a small set of characters coined to suit Western scientific notations. Graphic loans is a confusing term, since only a very small number of graphs really have been borrowed from other languages into Chinese (see 2.1.3). We argue that graphic loans are better subsumed under the term silent borrowing if only involving orthographic and semantic properties. However, those from Japanese, should be labeled Japanese loanwords and dealt with separately, as discussed in the final section.

---

[2] We use form, character and graph interchangeably if not noted otherwise.

[3] A Chinese character can express (a) a character-bound meaning due to its form, which is often the older original meaning; (b) a modern meaning due to its association with a sound; (c) and then the sound itself. Different graphs can have the same meaning. A single graph can have different meanings. This is true at the character level. The sub-components of a Chinese character work in a slightly different way. See footnote 36.

[4] We assume that the underlying principle of associating a thought to a symbol are inherently identical of a word and a numeral, or any symbol for that matter.

Table 1 Chinese words starting with a variable letter

Source	Chinese	Pinyin
x shares	x 股	x[1] gǔ
(N, X) rays	(N, X) 射線	([ən], [eɪ.kʰə.s])[2] shèxiàn
(X, Y) chromosomes	(X, Y) 染色體	([eɪ.kʰə.s], [uaɪ]) rǎnsètǐ
(X, Y, Z) axis	(X, Y, Z) 軸	([eɪ.kʰə.s], [uaɪ], [tsi]) zhóu
x V power supply	x V 電源	x ([vi]) diànyuán

[1] x stands for any letter or number

[2] the pronunciation of English letters as part of a Chinese word is given in IPA. IPA samples come from a Mainland Chinese native speaker living in Taiwan

2.1.1 Alphabetic words

There is an ongoing discussion as to the degree where alphabetic words count as loanwords, since native speakers of Chinese often create those words themselves (see Cook (2018); Ding et al. (2017)). Huang and Liu (2017) argue that the linguistic (syntactic) behavior of Mandarin alphabetic words (MAW) is similar to that of character words, even in light of rules for lexical word formation. In terms of pronunciation, however, alphabetic words are heavily influenced by the accent of the speaker (Ding et al. 2017). The IPA representations given in the tables below therefore serve as a reminder that they are not read with a standard English rules of pronunciation.

   (a) *Alphabetic words in Chinese starting with a variable letter or number* The Table 1 below lists words with a variable letter or number as part of the word. These are a subclass of Mandarin alphabetic words. Their number is much lower than words with a fixed alphabetic letter. Any such word can be represented in the format: $[x_a\mu_b] = [\langle\mu_a^{pc}\rangle_L\langle\mu_b^c\rangle_G]$, where the alphabetic component of the source word is kept in the transliteration $\langle\rangle_L$, or $\langle\rangle_S$ for any symbol, also preserving the sense and approximate pronunciation μ^{pc}. Source morpheme μ_b is mapped onto a Modern Standard Chinese morpheme written as a character $\langle\mu_b^c\rangle_G$, agreeing with its source counterpart only in semantic content, not in pronunciation (see Table 11:1).

   (b) *Words beginning with a fixed Latin and Greek letter* Words in this category contain a fixed alphabetic component as part of the word. Shi (2019) catalogs roughly 1,700 alphabetic words, most being purely alphabetic ones—abbreviations such as ITA, which stands for 'Information Technology Agreement', or words with a character component, such as *OK bianlidian* (OK 便利店 convenience store). Zhang et al. (2018) show significant differences in the selection and reception of Mandarin alphabetic words, listing such words as *byebye rou* (bye-bye 肉 flabby arms waving goodbye), or *band-zai* (band-仔 band member), *O-zui* (O-嘴 O-shaped mouth). Huang and Liu (2017) collect 56,833 MAWs combining the Sinica Corpus (Chen et al. 1996) and the Chinese Gigaword Corpus 2.0 (Huang 2009). In general, words beginning with a letter are more common than any other configuration (Huang and Liu 2017) (see Table 11:2).

(c) Words starting with a fixed number Shi (2019) finds 36 words starting with a fixed number. Huang and Liu (2017) do not collect this type of words. Any such word can be approximated as: $[x_a x_b m_c] = [\langle\mu_a^c\rangle_S \ \langle\sigma_b^p\rangle_L \ \langle\mu_c^c\rangle_G]$, whereas S represents any symbol, including numerals. Numbers are transliterated as morphemes, mapping onto their source counterparts only in terms of conceptual content *c*, but read in their native pronunciation, written in numeric form. The letters in 3C or 4K are kept as such $\langle\rangle_L$. Some numbered words are: *san* [si/ɕi] *chanye* (3C 產業) for '3C industry', *san* [ti] *dongzuo* (3 動作) for '3D action', *san* [ti] *xiaoguo* (3D 效果) for '3D effect', etc. (see Table 11:3).

2.1.2 Numerals, units and scientific notation

Arabic numerals from 1 to 9, as well as 0, can be considered loangraphs in Chinese (Xu and Zhang 2006, p. 206). Although Arabic numerals entered China much earlier during the 13th and 14th century, only one or two centuries after Fibonacci introduced Arabic numerals to Europe, their usage had not been widespread in China until the turn of the 20th century. Early translations of mathematical and scientific works used Arabic numerals, such as Xu Guangqi's (徐光啟) translation of Euclid's *Elements*, but his translation (and those of others) was mainly driven by the wish to promote the spread of Catholicism in China, and met with disinterest or rejection (Tian 2005, p. 57). Arabic numerals were first officially published in a mathematics textbook in 1892, translated by Calvin and Zou Liwen (鄒立文). Shortly before that, in 1885, a book for beginners in Western Calculation was published, only in Shanghainese (Xu and Zhang 2006, p. 75). Although the decimal system was a Chinese invention, the development of *wan* (萬 10,000) as a separate counting unit was probably an influence from translations of early Indian Buddhist scriptures, such as the *Avata.saka Sūtra* or *Flower Garland Sutra* (Li and Qian 1998, p. 230), and can maybe count as a loanword itself (see Table 11:4).

Mathematical symbols $(+, -, \times, \div, \%, =, x, \pi, <, \subset, \in, \infty, \int,$ etc.), units (*mm*, *cm*, *m*, *km*, m^2, m^3, *s*, *h*, etc.) and scientific notation in general can all be considered loangraphs. Every scientific notation is basically a word with a specific conceptual association and a pronunciation. The borrowing process of a scientific notation is similar to that of numerals: $[x] = [\langle\mu^c\rangle_S]$, where pronunciations mostly observe the native tradition, maybe with the exception of π and a few others (see again Table 11:4).

In Taiwan, the basic unit for measuring the size of living space is called *ping*[5] (坪), originally meant 'flat wide space', and is a loanword from Japanese. Characters such as *haoke* (毡, milligram), *qianwa* (瓩, kilowatt), or *jialun* (嗧, gallon), have all been created[6] at the beginning of the 20th century in response to Western units of

---

[5] One *ping* (or 'tsubo' in Japanese) = $3.3058 m^2$. This is not related to one *tatami*, which is $85 cm \times 180 cm$ or $1.53 m^2$. Introduced in Taiwan in 1895 and kept ever since then.

[6] Whether they have been invented in Japan is unclear yet. If so, they would count as Japanese loangraphs.

Gram and Watt. These can count as a special subgroup of loanwords, under which entirely new characters have been invented.[7]

2.1.3 Non-Chinese loangraphs

Arguably the best known loangraph in Chinese is the Buddhist symbol *wan* 卍, also written as 卐. Some controversy exists about the origin of the loangraph Pu (2001)[8] (see Table 11:5).

More recently, a couple of Japanese characters have gained recognition in the Chinese lexicon, e.g. the character *dan* (丼), originally meaning 'dwell', and mentioned in the *Shuowen* as a variant of *jing*丼 (dwell), despite its rare usage. *Dan* (丼) gained a new life, a new pronunciation (*dōng* or *dòng*) and a status of an independent character, differentiated from *jing* (丼), only after the Japanese used *dan* for referring to *donburi*, the name of a kind of Japanese food in a large rice bowl. With the global spread of Japanese cuisine, the character also gained popularity in Chinese communities. Another example is *xian* (腺 gland), etc. (see Table 11:6).

The borrowing of non-Chinese graphs as new graphs into the Chinese character set can be expressed as $[\langle m \rangle_g] = [\langle \mu^c \rangle_g]$. During the process, the borrowed graph *g* is linked with a native Chinese morpheme μ, i.e. its associated concept *c*. The orthographic form *g* is copied. Only rarely can it be found that a Japanese character is borrowed and combined with a Chinese ontological suffix, as in 鱈 $_{JP}$ → 鱈魚 $_{CN}$ (see Table 11:7).

2.1.4 Japanese loanwords

The special borrowing process between Japanese Kanji and Chinese characters has been labeled as 'graphic loans' (Masini 1997), or *jiexing hanzici* (借形漢字詞, cf. Shi (2013)), or *xingyici* (形譯詞, cf. Yang (2007)). In this paper we regard this process as 'silent borrowing' (see Table 11:8–10). The total number of Japanese loanwords in the Modern Chinese lexicon can figure between 1500, up to 3000 words (Shi 2019).

Strictly speaking, the term graphic loan is a misnomer. You cannot borrow what you already have in possession. The novelty of this type of borrowing process is the concept associated with the characters (which may be expressed by a novel combination of characters), but not the characters themselves (Wang 2015, p. 180). In case of any orthographic differences between Japanese and Chinese graphs, the Chinese do the exact opposite of borrowing, namely they replace Japanese graphs

---

[7] In China, these characters have been forced to retire in 1977 and substituted with their long version, e.g. 瓩 → 千瓦. They are still in use in Taiwan today.

[8] It is an often repeated story that Empress *Wu Zetian* decided the character to be pronounced *wan* and officially accepted it into the Chinese lexicon. Others argue, the character was already used on Chinese pottery during the Warring states (Shi et al. 1989). The matter is undecided.

with their own graphs, that is selecting the Chinese set-equivalent form, i.e. the 'variant'. This can be represented as: $[\langle m \rangle_g] = [\langle \mu^c \rangle_{g'}]$, whereas g' refers to the substitution of a character with its set-equivalent variant.[9] For someone unaware of the substitution, g' is equivalent to G.

What is meant with the terminology 'graphic loans', is the theoretical intention to differentiate between a silent borrowing process (borrowing meaning via the characters without the associated sound), and a material borrowing process that deals with sound and meaning.

Now, the importance of silent borrowing plays directly into the question of how to determine the status of Japanese loanwords. Two points are thus brought up for discussion: (a) how much emphasis is placed on sound for defining loanwords?—A steadily closing gap is clear in the academic discussion that recognizes the borrowing of word meaning as equally important as sound; (b) how many of those loanwords labeled 'Japanese' are of real Japanese origin?—a question more related to historical facts. Chen (2011); Feng (2016) argue that some Japanese loanwords might have Western origins. Generally speaking, the influence of Western (mostly) missionaries on some words in Chinese is a historical fact. Zhao (2016) shows for example that some mathematical terms like *xing* (形 form, shape), *jiao* (角 angle), *xian* (弦 arc) can be traced back to early translations from the *Elements of geometry* by Xu and Ricci, published in 1607. However, it has long been overlooked that the Western influence since the industrial revolution and the rise of the West is a much greater factor. Schmidt (*forthcoming*) argues that roughly a quarter of Japanese loanwords in Modern Chinese—as recognized by the current literature—are of Western origin[10]. Although oral borrowing and borrowing via informal writing can be fairly assumed to have happened, but the more traceable pathway of this transmission was by early English-Chinese dictionaries upon which Japanese translators heavily relied. Many of these dictionaries included new words (compounds, creations) for new concepts. One of those traceable impacts is Wilhelm Lobscheid's *English-Chinese Dictionary* (英華字典), published between 1866 and 1869, in which dozens of words made their first debut—words he and his Chinese co-lexicographers arguably created themselves (see Table 2).

---

[9] For example: Japanese 保$_{m_1}$ 險$_{m_2}$ → Chinese 保$_{\mu_1}$ 險$_{\mu_2}$, whereas m_2 is set-equivalent to μ_2. If we just look at Japanese loanwords in Chinese, according to our Japanese Loanword List—comprising currently 1468 at least double confirmed Japanese loanwords—roughly 146 types (characters) or 16% of total 911 character types are different. The most common ones are: Kanji (*pinyin* TC/SC): frequency. 学 (*xue* 學/学): 70; 働 (*dong* 動/动): 39; 体 (*ti* 體/体): 26; 権 (*quan* 權/权): 25; 会 (*hui* 會/会): 23; 国 (*guo* 國/国): 18; 観 (*guan* 觀/观): 16; 経 (*jing* 經/经): 13; 対 (*dui* 對/对): 12; 証 (*zheng* 證/证): 10; 点 (*dian* 點/点): 10.

[10] To verify the entrance of a Japanese loanword into the Japanese and Chinese lexicon, we rely on verification by dictionary, a process in two parts: (a) accessing two large digital databases, one specifically designed for early modern dictionaries, (i) the database for English Chinese dictionaries (Academia Sinica, Modern Historical Database), and (ii) the Chunagon (中納言, Balanced Corpus of Contemporary Written Japanese); and (b) checking all suggested Japanese loanwords against our own database.

Table 2 Selected words first attested in the English-Chinese Dictionary by Lobscheid

Source	Word	Pinyin
Bill	匯票	huìpiào
Civil law	民法	mínfǎ
Communication	交流	jiāoliú
Distribution	分配	fēnpèi
Effect	作用 [1]	zuòyòng
Exception	例外	lìwài
Human rights	人權	rénquán
Psychological	心理	xīnlǐ
Sovereignty	主權	zhǔquán
System	系統	xìtǒng

[1] 作用 meant 'great effort; to cast a spell' in Middle Chinese

In Japan, received and home-made loanwords were part of a large-scale social and linguistic modernization, strongly appealing to the Chinese people as well. The starting date for back borrowing from Japanese into Chinese can be set around the year 1900.[11] While it took nearly 30 years for the Japanese vocabulary to complete modernization, Chinese achieved its own 'second wave' of modernization[12] in roughly half the time.

2.2 Borrowing phonetic properties

Phonetic transliteration is the single uncontested category of loanword borrowing. The borrowing process of phonetic properties is understood in the following way: Ideally, each source morpheme can be mapped onto one or more target syllables in Chinese. That is, in a Chinese phonetic loanword, each syllable has no more and no less than a single phonetic function, referencing back to a source word. The actual process of phonological adjustment during phonetic borrowing may vary, according

---

[11] According to Qu (1973, p. 133) the first mission of thirteen self-paying Chinese exchange students to Japan was in March 1897, six followed in 1898, 18 students in 1899. From 1903 on, the then Qing government issued grants, consequently the number of students increased sharply into the thousands. On the other hand, the Ministry of Education in Japan (文部省) sent out 178 students to Western countries before 1899, but not a single one to China (Tan 2018).

[12] In general terms, it is debatable at what point in time the Chinese vocabulary can be considered modern since the written language was hardly reflective of the actual spoken vocabulary. However, in terms of modern loanwords, most scholars would draw a line at the beginning of the Opium wars (Feng 2004, pp. 15–23, Yang 2007, pp. 2–6, Fu 2011, p. 228). Robert Morrison published the first ever complete translation of the *New Testament* in 1814. Together with William Milne, he finished the translation of the Bible in 1823. His *Dictionary of the Chinese language* was published in 1822 (Foley 2009, pp. 18–19). So the consensus is to set the beginning of modern loanwords at around 1800. Yet, it is undoubtly the fact that a second, much stronger wave of modernization happened after the turn of the 19th century due to the influx of Japanese loanwords. We are referring here to this 'second wave' of modernization.

to various perception-based, phonetic or phonological rules (see Peperkamp et al. 2008; Paradis and Tremblay 2009; Kang 2011).

Prior to more discussion about simple phonetic loans in more detail, please note that Table 11 contains a rather large number of specific subcategories related to phonetic borrowing. Due to limited space, we cannot go into detail for all of these. We regard these types of words as phonetic loans because their most salient aspect of borrowing is based on phonetic representation. This main function can then be combined with other lexical processes. However, to establish a neat classification, we explicitly deny any morpheme-level semantic processes to participate in any borrowing process called 'phonetic'. In other words, as soon as sub-word semantic processes are combined with any phonetic processes, the entire categorization shifts towards complex transliteration. This being said, the classification does allow the interaction of orthography with phonetics in the following way: (a) complex orthography (*karaoke* 卡拉OK (see Table 11:12); also *asir*阿 'sir' (see Table 11:13); as well as 'byebye' → 88 (see Table 11:14); and (b) phonetic transliteration partly imitating source orthography, as in $[t^{h}i]xu$ T--恤 'T-shirt' (see Table 11:15). The mapping from a source syllable to target syllable during borrowing can be represented as $[s_a] = \langle \sigma_a^p \rangle$, using p if the target syllable approximately resembles the pronunciation of the source syllable.[13] When Chinese borrows polysyllabic source words, it is almost forced to apply a polysyllabic phonetic representation/transliteration, which motivates the language to adopt a more semantics-orientated transliteration strategy. In fact, phonetic borrowing is often only the first step of the borrowing process, followed by a collective effort from the language community to figure out better substitutes via either orthographic adjustment (i.e. by adding classifiers, prefixes, by shortening), or re-creating the word using some semantics-oriented approach (conceptual remake).

2.2.1 Material borrowing (simple phonetic loans)

Material borrowing refers to the borrowing of sound-meaning pairs, irrespective of orthographic properties (see Table 11:11). Specifically, material borrowing refers to word-level semantic borrowing only. If any morpheme-level semantic processes are involved, the category changes to complex transliteration. In the Western literature the range of material borrowing might include lexemes, or lexeme stems[14] (including also sometimes affixes[15]), to even entire phrases (Haspelmath and Tadmor 2009). However, the concept of material borrowing is complex in the

---

[13] A syllable without p would be a target syllable phonetically unrelated to the source word. We could also add a p to a morpheme (μ^p), which designates that the target morpheme is supposed to sound similar to the respective source subpart.

[14] In rare cases, phonetic loanwords are treated as stems combined with Chinese prefixes, such as *lao-* (老-), or *a-* (阿-); a rare example would be *lao K* (老 K 'King in a card game'), *laodou* (老豆 respectful son), and *laodike* (老狄克 honest person, gentleman).

[15] As a general rule, (a) loanwords almost never end in suffixes typical only for bound morphemes, such as *-er* (兒), *-tou* (頭), *-zi* (子). An exception could probably be *yuanzi* (原子- atom); and (b) loanwords only very seldom use reduplication of syllables, such as *momo* (嬷嬷), transliteration of Manchu *meme eniye*, or *niuniu* (妞妞) which comes from Manchu *nionio* 'little girl', allowed probably only because those terms are basically phonetic loans of kinship terms.

Chinese context, because phonological adjustment often results in polysyllabic constructions.[16] If such polysyllabic constructions, however, are supposed to be fully integrated into the native vocabulary, then they require any syllable to have a tone, which effectively turns the respective syllable into a morpheme. Once we start to deal with morphemes, we cross the border into complex transliteration territory. This does not mean that there cannot be material borrowings in Chinese, it means that because monosyllabic morphemes often take part in compounding polysyllabic loanwords, the function of each syllable is more than just syllabic. The morphemic quality of each unit plays an important part in supporting or influencing the entire word sense. Hence, it is often very difficult to say that each unit is solely restricted to a phonetic function, and only the entire phonetic form of the word is then equated with the entire word sense. If so, then this is most likely to happen with proper names. Finally, in the absence of morphology, compounding plays a very dominant role in Chinese, which always involves the semantics at the morpheme-level. Yet, sub-word level compounding processes are generally excluded from loanword theories, a mistake any Chinese loanword theory should carefully avoid.

Simple phonetic loans in Chinese are easily noticeable for a native speaker because of their phonetic resemblance to the foreign source words, and/or, in terms of their written form, due to their apparent lack of a semantic relationship with the chosen characters. In light of a multiplicity of Chinese homonyms, the criteria of unanalyzability—as stressed by Haspelmath (2009)—can be especially useful on the orthographic level, rather than on the sound level.[17] This can be represented as: $[m_a m_b] = \langle \sigma_a^p \rangle_G \langle \sigma_b^p \rangle_G$.[18]

Generally speaking, in any given scenario of linguistic contact between Chinese and a language without script or with a non-Han-related script, it would be reasonable to expect the way of vocabulary transmission to be mainly phonetic. This certainly is true throughout history, although the linguistic footprint for various Mongolian tribes was astonishingly limited, e.g. phonetic loanwords from the Xianbei (鮮卑) during the Late Han dynasty, such as *kehan* < *qaghan* (可汗 Supreme leader), or the proto-Mongolian Khitan (契丹) at the times of Northern Liao, Eastern Liao and Later Liao, e.g. *mouke* < *mäke* (謀克 organization of Khitan families), and the Tungusic-speaking Jurchen (女真) during the Jin dynasty, e.g. the name of the *Jurchen* itself stemming from *nüzhen* < *jurchin*.

---

[16] However, a few monosyllabic phonetic loanwords do exist in Chinese, e.g. *ta* (塔 tower), *fo* (佛 Buddha), *beng* (泵 pump), *an* (氨 ammonia), *dian* (碘 iodine), *ka* (卡 card), even alphabetic words, such as 'A' (to steal) as in 'A 走', see (Cook 2018, p.14). However, the origin of 'A' is debatable. These all operate as free morphemes. Furthermore, even disyllabic but monomorphic phonetic loanwords can be found already in much older strata of the Chinese vocabulary (Xu (2013)), such as *furong* (芙蓉 hibiscus, lotus flower), *huangtang* (荒唐 absurd, ridiculous), *hupo* (琥珀 amber), *kafei* (咖啡 coffee), *lese/laji* (垃圾 trash), *manao* (瑪瑙 agates, cornelian), *moli* (茉莉 jasmine), *pipa* (枇杷 loquat), *pili* (霹靂 thunderbolt), *popo* (婆婆 mother-in-law), *putao* (葡萄 grape), *shanhu* (珊瑚 coral), *suanni* (狻猊 lion, older word for *shizi* 'lion', *tuoluo* (陀螺 spinning top toy), etc. (For reduplication of morphemes see Booij and Van Marle (2001, p. 290–292), whether disyllabic words are stems, see Baxter and Sagart (2014, p. 391).)

[17] During audible perception, the syllable combination cannot be resolved into a cluster of morphemes.

[18] We do not write $\langle \sigma_b^p \rangle_G^c$ in order to stress the lack of a conceptual relationship between the characters, although every character obviously has a meaning, too.

During the Mongolian-reigned Yuan dynasty, Chinese received a few phonetic loans such as *chengjisihan* < *čingizhan* (成吉思汗 Genghis Khan), *gebi* < *gobi* (戈壁 desert), *hutong* < *gudum* (胡同、衚衕 alley), *laba* < *labai* (喇叭 trumpet), or *haba (gou)* < *xaba, halban* (哈叭狗 (hairy lion) dog), etc.[19] Despite the 300 years of lexical impact accompanying Manchu rule (Qing dynasty) over China[20], only a couple of words were left behind, such as *bashi* < *baksi, bagši* (把勢 accomplished technician, master)[21], *saqima* < *sacima* (薩其馬 Sachima, a common Chinese pastry), *mahu* < *mahuu* (馬虎 Manchu hood, wrapped around the ears and forehead, but open on the top), etc.

This demonstrates that even phonetic loanword transmission from source to target languages is to some degree facilitated by some sort of cultural influence of the donor over the target culture, whereas a lack thereof results in a low number of loanwords. This explains why phonetic loanwords from Sanskrit and Indian culture are comparatively more common in number and usage, e.g. *emituofo* < *amitābha* (阿彌陀佛), *luohan* < *arhat, arhān* (羅漢 Rohan), *fan* < *brahmā, brahman* (梵 clean, Sanskrit), *fo* < *Buddha* (佛), *zhina* < *čina* (支那 China), *heshang* < *khosha* (和尚 teacher, monk), *yujia* < *yoga* (瑜伽), etc.

Phonetic borrowing has also been the standard approach by Yan Fu and other early Chinese foreign-trained scholars (see Table 11:16). 'Zebra' was phonetically translated into *zhibula* (today semantically *banma* 'striped-horse'), or 'philosophy' into *feiluosufei* (today *zhexue* 'wise-learning'), or 'revolution' as *lifoliuxuan* (today *geming* 'change-heavenly mandate').[22]

A couple of reasons made this approach difficult. Firstly, words of modern science often have multiple sources: English, French, German, Russian, etc. A phonetic representation makes back-tracking from sound to meaning difficult. Secondly, phonetic transliteration often requires the use of multiple syllables to convey a single concept, which is unnatural in Chinese. Moreover, many of those words use difficult characters to stress their phonetic function (within that word) in contrast to a literal semantic reading. This, however, makes the word harder to write and remember. Lastly, using complicated polysyllabic words in Chinese might be acceptable in specific religious texts, e.g. in translations related to Buddhism. However, it would be rather discouraging to use obscure characters and long

---

[19] https://languagelog.ldc.upenn.edu/nll/?p=31837.

[20] Compare this to the impact French had upon the English vocabulary during a comparable 300 years of occupation of England by the Normans.

[21] *bashi* (把勢) is itself back imported from Chinese *boshi* (博士 knowledgeable man).

[22] 'Revolution' *geming* (革命) originally meant 'change (革) of heavenly mandate (命)', referring to a change from one dynasty to another, and was not listed in English-Chinese dictionaries as a translation for the word 'revolution' until around the turn of the 20th century. The English-Chinese Lexicon by Guang Qizhao (鄺其照) translates *revolution* as 'period' (*yizhou* 一周) in 1868 and as 'rotation' (*xuanzhuan yihui* 旋轉一回) in 1875; Shi (2019) says *geming* (革命) appeared with its modern sense of 'violent overthrow of the government' first in Chinese in 1896 in *Shiwubao* (時務報, 1896-12-05).

multisyllabic words[23] for new everyday-concepts such as 'telephone' (*delüfeng* 德律風) or 'president' (*bolixitiande* 伯理璽天德) etc.

Early simple phonetic transliterations might be adjusted over time. Indeed, the most common used fall under the categories of (a) orthographic adjustment, not changing pronunciation, e.g. 'cacao' *keke* 柯柯 → *keke* 可可, or 'curry' *gali* 咖喱[24] →咖哩, or 'Italy' *yidali* 義大利 (as used in Taiwan) →意大利 (as used in Mainland China, see Table 11:17); (b) shortening, with optional addition of ontological suffix, e.g. 'America, USA' in *meilijian* 美利堅 → *mei* + 'country' (美 + 國); (c) replacement by semantic transliteration, e.g. professor $_a$ from its original phonetic loan *bufeise* (卜 $_{a_1}$ 非 $_{a_2}$ 色 $_{a_3}$ foretell$_{a_1}$ not$_{a_2}$ color$_{a_3}$) to later re-coined *jiaoshou* (教$_{a_1}$ 授$_{a_2}$ teach$_{a_1}$ give$_{a_2}$). For more examples, see Zeng (2017, p. 130–132).

2.2.2 Mediated material borrowing

Mediated material borrowing refers to a scenario where a target language borrows a sound-meaning pair via an intermediate third-party language. There can be more than one intermediate language (see Table 11:18). For example, 'Aspirin' was materially borrowed from German to English, then to Japanese (*asupirin*), and lastly into Chinese (*asipilin*). Japanese loanwords, written in Katakana or Hiragana, serve as mediator for materially borrowing into Chinese (see Table 3). If Japanese words are written in Kanji, Chinese speakers would prefer to borrow the orthographic units (characters) over sound, resulting in graph-meaning borrowing, which is labeled herein as silent borrowing, or otherwise as graphic loans (Tranter 2009), and symbolic loans (Cook 2018).

In some cases, the borrowing from Japanese into Modern Standard Chinese is mediated by a dialect of Mandarin, e.g. Southern Min, as this is sometimes the case in Taiwan, for example *tenpura*[25] (天ぷら) was mediated by Southern Min thian$_{35}$ pu$_{55}$ lah$_3$ and then adopted further into Chinese *tianbula* (甜不辣) (see Table 11:19).

2.3 Borrowing semantic properties

Discussing loanwords in Chinese, we often see a differentiation between phonetic transliteration (*yinyi* 音譯) and semantic transliteration (*yiyi* 意譯), as in Zhao (2016, p. 3) However, what counts as a semantic transliteration though has not been strictly defined. For example, the words *miyue* (蜜月 honeymoon), *bairimeng* (白日夢 daydream), and *dianhua* (電話 telephone) are all considered semantic loanwords of some sort in the current literature, either implicitly as *jiexing hanzici*

---

[23] In phonetic transliterations the sense of Chinese characters play a far lesser role, however, in reality the inherent sense of a character cannot be completely suppressed even in phonetic readings, which might cause distraction. Note that the very first transliteration of 'Plato' was *baladu* 霸$_a$ 辣$_b$ 篤$_c$, which literally meant: bully $_a$ spicy $_b$ earnest $_c$, and was later re-transliterated into *bolatu* 柏$_a$ 拉$_b$ 圖$_c$, which less dramatically reads now as: Cypress $_a$ draw $_b$ picture $_c$.

[24] Early variants are 加里 jiālǐ, 加釐 jiālí, 咖唎 gālì/kālì.

[25] Also phonetically romanized as *tempura*.

Table 3 Examples of loanwords into Taiwanese Mandarin from English via Japanese by material borrowing

Source	Japanese	Romaji$_{\text{JP}}$	Mandarin$_{\text{TW}}$	Pinyin
Aspirin	アスピリン	Asupirin	阿斯匹林	āsīpīlín
Bus	バス	Basu	巴士	bāshì
Camera	カメラ	Kamera	開麥拉 [1]	kāimàilā
Engine	エンジン	Enjin	引擎 [2]	yǐnqíng
Percent	パーセント	Paasento	趴 [3]	pā

[1] In Southern Min kha$_{33}$ me$_{55}$ lah$_3$

[2] In Southern Min ian$_{35}$ jin$_{51}$

[3] In Southern Min pha$_{35}$ sian$_{51}$ too$_{11}$

(借形漢字詞, cf. Shi (2013))[26], or explicitly as *yiyi* (意譯, cf. Yang (2007)), *duiyi* (對譯, cf. Feng 2004, p. 27), or *hezhihanyu* (和製漢語, cf. Takeyoshi 1996). However, if we take a closer look at the morpheme level, we can see a different type of relationship between the source-morpheme structures and target-morpheme structures for each of the above-mentioned examples. In the case of honey$_a$ moon$_b$, each morpheme of the source word is represented in the target word 蜜 $_a$ 月 $_b$, while also preserving its morphosyntactic order. Due to said property, we call this type of transliteration iso-morphemic.[27]

In the second case, day$_a$dream$_b$ has two morphemes, but its transliteration *bairimeng* 白 $_{\mu_1}$ 日 $_{\mu_2}$ 夢 $_b$ has three (μ_1, μ_2, b), while only morpheme b is directly projected onto the target word. The first morpheme day $_a$ is split into a 'two-syllable-two-morpheme'-construction, roughly meaning 'white $_{\mu_1}$ day $_{\mu_2}$'.[28] Due to this mismatch in the morphemic structure – based either on a different count of total morphemes, a partial morphemic mapping, or the addition of ontological suffixes to the target word, we call this type complex transliteration.

The third common type of transliteration is a conceptual remaking, most commonly observed in Japanese loanwords, as in the case of tele $_a$phone$_b$, which, if using an iso-morphemic projection, would be transliterated as *yuansheng *遠 $_a$ 聲 $_b$ for *tele* = far$_a$ and *phone* = sound$_b$. However, the modern Chinese version is *dianhua* 電$_{\mu_1}$ 話$_{\mu_2}$, which reads as 'electric$_{\mu_1}$ talk$_{\mu_2}$'. The modern word is connected to the source word neither by morpheme structure nor by phonetic properties; it is an brand-new word created based on the shared concepts of both words, namely what a telephone is (Table 4).

---

[26] Graphic loaning is considered silent borrowing, borrowing the sense via the graph, but not their pronunciation.

[27] Preserving order is not a necessary condition, important is,however, that all morphemes are projected in total, or nothing else is added.

[28] There would have been other options available for 'day' in the Chinese lexicon, for example: *tian* 天, *ri* 日, or *baitian* 白天.

Table 4 Comparison of transliterations

Source	Target		
	Iso-morphemic	Complex	Conceptual
Honey$_a$ moon$_b$	蜜$_a$ 月$_b$ [1]	-	?婚後假期
Ultra$_a$ violet$_b$ rays$_c$	紫$_b$ 外$_a$ 線$_c$	-	-
Day$_a$ dream$_b$	-	白$_{\mu_1}$ 日$_{\mu_2}$ 夢$_b$ [2]	-
Plastic$_a$ box$_b$	-	塑$_{a_1}$ 膠$_{a_2}$ 盒$_b$ [3]	-
Tele$_a$phone$_b$	*遠$_a$ 聲$_b$	-	電$_{\mu_1}$ 話$_{\mu2}$
Rain$_a$bow$_b$	*雨$_a$ 弓$_b$	-	彩$_{\mu_1}$虹$_{\mu_2}$

[1] 婚後假期 hūn-hòu-jiàqí [marriage-after-holidays] [2] 白 日báirì is not a free morpheme combination in modern Chinese, or in other words, it is not a modern word, but used nowadays only in fixed combinations such as 白日鬼 báirìguǐ (strange person, 'ghost during daytime'), 白日撞 báirìzhuàng (robbery, 'accident during daytime'), 白日賊 báirìzéi (robber, 'thief during day time'), 白日衣綉 báirìyīxiù (daily clothes), etc. [3] 'plastic' is one morpheme in English but two morphemes in Chinese: 塑$_{\mu_1}$ 膠$_{\mu_2}$, sùjiāo whereas μ_1 = to form, originally with clay, to make porcelain products, \mu _2 = glue, adhesive. Both morphemes combined express the concept 'plastic'

2.3.1 Iso-morphemic transliterations

An iso-morphemic transliteration is a near-synonymous mapping relationship between morphemes of the source and target words. Morphosyntactic order is often preserved. Since strict synonyms are rare, some degree of sense variation should be allowed, for example, German 'Blind$_a$darm$_b$' (English: 'blind$_a$ gut$_b$', appendix, caecum), was the source for Chinese *mangchang* 盲$_a$ 腸$_b$; however there is arguably a slight variation in the scope of the senses Darm$_b$ and *chang* (腸$_b$), because the Chinese pendent can also refer to 'thought' and 'feeling'. However, since the core senses of both morphemes are near-synonyms, we can still consider this transliteration to be iso-morphemic. The following list shows some similar examples of loanwords, patterned in an iso-morphemic fashion. This can be written in a general form as: $[m_a m_b] = [\langle \mu_a^c \rangle_G \langle \mu_b^c \rangle_G]$, whereas each morpheme has an external direct conceptual mapping relationship with its source counterpart $\langle x^c \rangle$ (see Table 11:20) (Table 5).

2.3.2 Complex transliterations

Complex transliteration refers to an imperfect or partial mapping relationship between the morpheme structure of a source word and a target word. During borrowing, as long as phonetic and semantic processes are involved at the morpheme-level, or if any division of labor takes place between a syllable and a semantically relevant graph, then we have a case of complex transliteration. While material borrowing only refers to a semantic mapping relationship between source and target word at the word level, in complex transliterations the morphemes of the target language play a much bigger role. They attempt to exploit the vast number of

Table 5 List of loanwords by Direct Iso-morphemic transliteration

Source	Chinese	Pinyin	Morpheme Sense
Dumb$_a$ bell$_b$	啞$_a$ 鈴$_b$	yǎlíng	dumb$_a$ bell$_b$
Gastric$_a$ juistice$_b$	胃$_a$ 液$_b$	wèiyè	stomach$_a$ liquid$_b$
Hartes$_a$ Wasser$_b$	硬$_a$ 水$_b$	yìngshuǐ	hard$_a$ water$_b$
Hot$_a$ dog$_b$	熱$_a$ 狗$_b$	règǒu	hot$_a$ dog$_b$
Most$_a$ favourite$_b$ nation$_c$	最$_a$ 惠$_b$ 國$_c$	zuìhuìguó	most$_a$ favorite$_b$ nation$_c$

Chinese homonyms, but also interplay with the semantic functions of Chinese characters, often purposefully create a witty second layer of meaning. They do not intend to be exact transliterations. This part of the discussion is hardly touched upon outside of the Chinese context.

Now, the mismatch between the source and target morpheme structures can result from several reasons: (a) A set of functionally separated morphemes used in the same word, as in *jiuba* (酒吧 'bar'), here *jiu* ('wine, alcoholic') has a purely semantic function, *ba* a purely phonetic one. Or as in *kache* (卡車 '*orig.* car; truck', see Table 11:21). (b) A partial projection of the morphemes, as in Consul$_a$ general$_b$ transforming to *zonglingshi* (總$_b$ 領$_{\mu_1}$ 事$_{\mu_2}$). Here, only the morpheme 總$_b$ maps onto general$_b$ while 領$_{\mu_1}$ 事$_{\mu_2}$ is a conceptual remake of Consul$_a$, because both words are not related to each other in light of morpheme structure or phonetic properties. The same is true for *motuoche* (摩托車 'motorcycle'), here one part of the source word ('cycle') was substituted for a better fitting morpheme (*che* 'vehicle'), creating a mismatch on the morpheme level (see Table 11:22). (c) A partial projection with an addition of affixes in the target word, mostly loan-affixes but also autochthonous affixes. For example, 'assimilation' is projected into Chinese as similar$_a$ change$_b$, or *tonghua* (同$_a$ 化$_b$). The core sense of assimilation is represented by the morpheme *tong* (同$_a$), and the second morpheme *hua* (化$_b$) is a loan-suffix – a way of translating the English suffix *-tion* into Chinese. Another example would be *moshui* (墨水) for *ink*, where *shui* 'water, fluid' is the added suffix (see Table 11:23). In rare cases, a prefix can be used which normally only goes together with native Chinese morphemes, such as *a* in *ashe* (裕祥 'sir'), whereas the *a* is an honorary prefix (see Table 11:24). (d) A fusion of phonetic and semantic function for all morphemes (fully fused). Here, the complex transliteration can be so successful that it is hard to spot: *jiapan* < 阿蛇 'chapan' from the Turkish *chapgan*, meaning 'sewn together', whereas the pronunciation, the word meaning and the sense of each individual morpheme are all reasonably well mapped to the source word (see Table 11:25). (e) A fusion of phonetic and semantic functions for some units of the target word (partially fused). 'Idol $_a$' is projected to *aidou* 愛$_{\mu_1}$ 豆$_{\mu_2}$, where μ_1 and μ_2 both imitate the pronunciation of the source word, but

μ_1 also carries the sense of 'love', to express the positive feelings one has towards an idol. *Shuqiu* (述$_{\sigma_1}$ 球$_{\mu_2}$) for 'shoot' would be another good example. Here, σ_1 is only phonetic, but μ_2 is phonetic and somehow loosely semantic (see Table 11:27).

An important property of many but not all complex transliterations is the mismatch of the unit (morpheme or syllable) count, generalized as $[m_a] = [\langle\mu_a\rangle_\mathrm{G} \langle\mu\rangle_\mathrm{G}]$, whereas the difference of units is at least one, either one more or one less. For example, 'gene$_a$' has one morpheme, but its Chinese equivalent *jiyin* (基$_{a_1}$ 因$_{a_2}$) counts two; 'sauna$_a$' counts three in *sanwennuan* (三$_{a_1}$ 溫$_{a_2}$ 暖$_{a_3}$).

Given that this type of borrowing process has been widely gone unacknowledged in the literature, we think it is useful to discuss it here in slightly more detail:

(a) Complex transliteration with functionally separated morphemes The simplest situation of complex transliteration happens if all morphemes of the target word are functionally separated. During this process, the target language opts for using different morphemes establishing either a phonetic or a semantic relationship to the source word.

'Ballet' is transliterated into *baleiwu* (芭蕾舞 bā$_{a_1}^p$ lěi$_{a_2}^p$ dance$_b^c$), whereas *balei* has a phonetic function and *wu* a semantic function; or *lace$_a$ border$_b$* into *leisi huabian* (蕾絲花邊 lěi$_{a_1}$ sī$_{a_2}$ flower$_\mu$ border$_b$), whereas *leisi* is phonetic and *huabian* semantic. In some cases, this works even for abbreviations, such as 'AIDS', here transliterated into *aizibing* (艾滋病[29] ài$_{a_1}^p$ zī$_{a_2}^p$ sickness$_b^c$), whereas *aizi* refers to AIDS phonetically, and *bing* ('sickness') semantically.[30]

(b) Complex transliteration with suffixes During transliteration, Chinese lexical word-building rules allow the addition of either ontological suffixes, such as *-che* (∼車 car), *-fang* (∼房 house), or the addition of more abstract suffixes, most of which are loans themselves from Japanese, such as *xing* (∼性-tion), or *-shi* (∼式 kind of, style), etc. Here, it is important to note that the sense of those suffixes have undergone some significant semantic shift (cf. Table 8)[31] (see Table 11:23 and 24) (Table 6).

(c) Complex transliteration with functionally fused morphemes A more complicated situation surfaces, if a single target morpheme has more than one function, most often a fusion of phonetic and semantic functions. The fusion of phonetic and semantic functions happens in two different directions. Firstly, a syllable of the source word is phonetically related to one or more target morphemes

---

[29] In Taiwan: *aizibing* (愛滋病). In the Taiwanese version of the loanword, the first syllable is represented by a semantically relevant graph *ai* (愛 'love'), which allows for some kind of double reading, effectively changing the borrowing processes from functionally seperated morphemes to functionally fused morphemes.

[30] Strictly speaking the semantic relationship is established on grounds of AIDS bing a disease, and not specifically to the meaning of the acronym 'Acquired Immune Deficiency Syndrome'.

[31] For example, *lun* (∼論) originally meant 'convict, accuse, say, according' in Middle Chinese, and developed towards 'argue, debate, conclude'. *Lilun* (理論) was mentioned by Medhurst in 1848 as 'argue, debate' not 'theory', and then by Lobscheid in 1869 as 'discuss, discourse, reason, logic', but not 'theory'. In 1900, Liu 1900 (2018) mentions *lun* 論 in his schoolbook as 'conclude, final evaluation'; Shi (2019) dates the earliest attest in Chinese as 'theory' to 1894.

Table 6 Examples of complex transliteration with classifiers

Source	Chinese	Semantic transliteration
Empiric$_a$ism$_b$	經$_{a_1}$ 驗$_{a_2}$ 論$_b$	Experience$_{a_1-a_2}$ theory$_b$
Harden$_a$ing$_b$	硬$_a$化$_b$	Hard$_a$ change$_b$
Precis$_a$ion$_b$	精$_a$度$_b$	Fine$_a$ degree$_b$
Social$_a$ism$_b$	社$_{a_1}$會$_{a_2}$主$_{b_1}$義$_{b_2}$	Society$_a$ doctrine$_b$
Vernunft$_a$	理$_a$性$_\mu$	Reason$_a$ essence/nature$_\mu$

$[s_a] = [\mu^p_{a_1}\mu^p_{a_2}]$. Secondly, a morpheme (or shadow morpheme[32]) of the target word is semantically related to other morphemes (or shadow morphemes) of the target word $[\mu^{c'}_a \mu^{c'}_b]$, creating an additional word-internal layer of meaning, engaging in a conceptual reinterpretation (c') of the source morpheme $[m]^c$.[33] Fusion can take place fully, as well as fully based on an abbreviated source word, or partly. An example for partial fusion would be *miniqun* for 'mini$_a$skirt$_b$' (迷$_{a_1}$你$_{a_2}$裙$_b$), here the first two syllables are used phonetically, yet the characters used to represent those syllables create a second-layer meaning, roughly something like 'fan of you' or 'follow you', which obviously is not an intended transliteration of 'mini' and also not a conceptual remake, rather a play of meaning, given the constraints that the syllables are supposed to primarily represent the sound 'mini'.

2.3.3 Conceptual remakes

Conceptual remaking comprises the third subclass of semantic transliterations. These words are most detached from a direct morpheme-to-morpheme transliteration. In fact, conceptual remakes arguably are better described as words constructed in such a way that their morphemes capture the same idea of the source word in a completely new fashion, nevertheless referring to the same idea (see Table 11:30). 'Baseball' is a word named for some sort of ball game. The word highlights the importance of 'base' and 'ball' as two key features of the game. In Chinese, however, the word *bangqiu* (棒$_a$ 球$_b$) is used. It highlights bat$_a$ and ball$_b$ instead.[34] Using 'batball' instead of 'baseball' is a conceptual remake of the same idea of that very ball game. We can compare both approaches more generally in the following way: [base$_a$ ball$_b$] = [bat$_\mu$

---

[32] A shadow morpheme is a subpart of a word with fused phonetic and semantic functions from two different levels of analaysis. On the first level, the syllable is used for its phonetic relation to a source syllable. However, on the level of writing, the semantic function of its graph provides a second layer of semantic content. Shadow morphemes are interpreted as if they were morphemes by virtue of their respective graph senses. This is represented as $[\langle\sigma^p_{a_1}\rangle^{c'}_G\langle\sigma^p_{a_2}\rangle^{c'}_G]^{c'}$, whereas the indicator of the conceptual function, superscript c, is primed (c') to show that the meaning construed by the reading of the shadow morphemes $\langle\sigma_{a_1}\rangle\langle\sigma_{a_2}\rangle$ is a conceptual remake compared to the entire word $[\mu_a]$. Furthermore, please note that c is placed on top of G to show that the sense is related to the graph, not the morpheme.

[33] This happens of course on top of the standard word-to-word linkage, in simplified form $[m_a]^c = [\mu_a]^c$, whereas the word-sense of m is linked to the entire target word form μ.

[34] Of course, the morpheme *bàng*$_a$ has the nice feature of not only being conceptually related to the game, but also additionally having a pronunciation somehow close to the sound when the ball hits the bat.

Table 7 Examples of conceptual remakes

Source	Target	Morpheme sense	Notation
Tele$_a$phone$_b$	電$_a$ 話$_b$	Electric$_a$ talk$_b$	$[\langle\mu_a\rangle_\mathrm{G}\langle\mu_b^{c'}\rangle_\mathrm{G}]^{c'}$
Finance$_a$	金$_{a_1}$ 融$_{a_2}$	Gold$_{a_1}$ dissolve$_{a_2}$	$[\langle\mu_{a_1}^{c'}\rangle_\mathrm{G}\langle\mu_{a_2}\rangle_\mathrm{G}]^{c'}$
Nerve$_a$	神$_{a_1}$ 經$_{a_2}$	God$_{a_1}$ top-down relation$_{a_2}$	$[\langle\mu_{a_1}\rangle_\mathrm{G}\langle\mu_{a_2}\rangle_\mathrm{G}]^{c'}$
Objective$_a$	客$_{a_1}$ 觀$_{a_2}$	Guest$_{a_1}$ view$_{a_2}$	$[\langle\mu_{a_1}\rangle_\mathrm{G}\langle\mu_{a_2}\rangle_\mathrm{G}]^{c'}$
Personality$_a$	個$_{a_1}$ 性$_{a_2}$	Individual$_{a_1}$ essence / nature$_{a_2}$	$[\langle\mu_{a_1}\rangle_\mathrm{G}\langle\mu_{a_2}^{c'}\rangle_\mathrm{G}]^{c'}$

ball$_b^c]^{c'}$, where source morpheme base$_a$ is substituted with morpheme bat$_\mu$ in Chinese; ball$_b$ remains in the transliteration; the new word denotes the same concept by highlighting one different and one identical aspect of its original. The fact that bat$_\mu$ and ball$_b^c$ create a conceptual variation of the same idea is captured by superscript c', which stands for 'conceptual remake'. Yet, not all conceptual remakes are loanwords. As pointed out in Cook (2018), the Western literature deals with conceptual remakes often as examples of compounding. She suggests labeling conceptual remakes (in her terminology called 'holistic calques') as creations and not as loanwords. Although we agree that modern autochthonous creations are common, it is also true that many of the words that would fit into the category of 'holistic calques' are factually borrowed, which can be verified by comparing historic (bilingual) dictionaries of English, Chinese and Japanese. A categorization can not simply solve the problem based on some lexical principle (Table 7).

2.3.4 Sense extensions

Sense extension is the semantic process by which a target morpheme acquires an additional sense via borrowing without losing its original sense(s). Such an 'enhanced' suffix is then combined with standard morphemes to form new compounds. The enhancement stems from a period of social and linguistic modernization in Japan. Many of these words can already be found in classic Chinese texts, but not with their new meanings. The sense extensions are often semantically natural extensions of their original meanings. Note that many suffixes are highly productive (see Table 11:31).

The following list shows examples of suffixes which have undergone some sort of semantic extension by the Japanese, compared to their Chinese counterparts. Their number far exceeds one hundred. For example, *bao* (報) originally meant 'execute, answer, report' during the Han dynasty. The word was then borrowed into Japanese and acquired there, much later though, the sense 'newspaper'. After back-import into Chinese in the early 20th century, this new sense of 'newspaper' was added to the original sense 'answer, report', effectively creating a loan meaning extension. Today we find compound words such as *huibao* (回 報 'answer') with the old meaning, next to *zhoubao* (週 報 'weekly paper') with the new meaning (Table 8).

Table 8 Selected loan suffixes from Japanese

Suffix	Pinyin	Semantic Extension	Japanese loanwords in Chinese
～案	àn	*orig.* plate, table, to press, verify; + **case, matter**	方案、提案、預算案、圖案、議案
～觀	guān	*orig.* inspect, observe; + **way of seeing and thinking, -ism**	悲觀、參觀、界觀、客觀、樂觀、人生觀、唯物史觀、宇宙觀、主觀
～化	huà	*orig.* change, creation, education; + **-tion**	長期化、工業化、進化、口語化、綠化、強化、弱化、現代化、自動化
～覺	jué	*orig.* understand, discover; + **sense, feeling**	自覺、味覺、直覺、感覺、錯覺
～論	lùn	*orig.* discussion, verdict; *(MC)* report sb., speak; + **theory, summary**	辯論、法學通論、方法論、概論、結論、認識論、唯神論、相對論
～權	quán	*orig.* scales, might, power; *(MC)* substitute; + **right**	人權、主權、所有權、版權、財產權、參政權、債權、選舉權、霸權
～式	shì	*orig.* standard, use; + **style, method, -ish**	方程式、方式、告別式、恆等式、金婚式、流動式、形式、樣式
～問題	wèntí	*orig.* question; + **problem**	婦人問題、國際問題、民族問題
～性	xìng	expressing a property, ability, abstract concept, used for translating words ending in **-ity**	必然性、創造性、蓋然性、個性、可能性、偶然性、人間性、屬性
～主義	zhǔyì	*orig.* + **-ism**	保守主義、帝國主義、共產主義

2.4 Development of Chinese loanword theories

Generally speaking, phonetic borrowing, especially its most prototypical cases, such as material borrowing and phonetic transliteration of names, are academically accepted as criteria for loanwords across the board. Complex transliterations with fused or separated phonetic and semantic functions are also relatively undisputed, because phonetic features in the target word are partially reflective of the source language, either being phonetic (loanblend) in nature but adding a conceptual suffix, or due to the selection of graphs that express phonetic and semantic features simultaneously or separately. Calques are more often disputed, because they focus on the semantic representation of the source morphemes and are therefore less or not at all reflective of the phonetic properties of the source word.

The real challenge for any theory of loanwords in Modern Standard Chinese is when it comes to Japanese loanwords, because they predominantly are conceptual remakes and sense extensions. The following discussion focuses on the theoretical implications of identifying Japanese loanwords and grating them loanword status.

First of all, Japanese loanwords are highly integrated into the Chinese lexicon and syntactically not at all obviously different from any other words of Chinese origin. The dispute can be simplified down to two points: (a) the creatorship of the writing system and the characters pose the question of who came first and what matters most: the new words (semantic content, conceptual associations) that have been created in Japan, or the characters (orthographic medium) by which those new words are recorded in writing, which are of Chinese origin; and (b) the peculiar properties of the Chinese characters to carry phonetic and semantic functions, which can be independent of each other activated or read.

In the European tradition, the concept of 'loanword' was developed based on languages written in some sort of alphabet. Being alphabetic/phonetic in nature, a word can be judged as loanword by directly observing its phonetic and morphological properties, i.e. whether it is a 'phonological loan' (Haugen 1972), also called 'transfer' (Weinreich 1953), or whether only the meaning has been borrowed in some way, called 'loanshift' (Haugen 1972), further differentiated by Weinreich (1953), whether all morphemes were borrowed, called 'transliteration loan', or just a single morpheme, called 'semantic loan'. Following the Western tradition, Wang (1958) and others (Liu 1984; Lü 1992; Li and Yu 2005), propose a strict loanword definition:

Strict Definition A word is granted loanword status if, and only if, both of the following conditions are true:

- the word is not analyzable
- the word is pronounced resembling the source language.

The same logic is applied by many Chinese scholars of Japanese loanwords (Zhang 1958; Wang 1980; Liu et al. 1984; Zhou 1995). For instance, (Wang 1980) argues that '*qudi*, *yindu*, *shouxu*... do not really count as loanwords, because their sound has not been borrowed,' followed by Zhou (1995) differentiating sense extensions of existing words and new words created by adhering to Chinese word formation rules. Many Western scholars follow some flavor of the strict phonological definition, unwilling to grant Japanese loanwords full loanword status, labeling them 'graphic loans' instead (Masini 1997, p. 153; Tranter (2009); Wiebusch and Tadmor 2009, p. 581).

Theoretical advancements were possible by focusing more on borrowing sense than sound. Consequently, the restriction of phonetic borrowing as a necessary condition of granting loanword status, was softened (Zhou 1998; Shi 2000, 2019; Yang 2007; Xue 2007). Shi (2000) acknowledges this, stressing form and authorship over phonetic properties: "Japanese loanwords in the Chinese language need to be given a special status. One can say, these words are positioned between Chinese words and real loanwords. They are Japanese-made words, ... they borrow the word form (i.e. characters), but not the pronunciation, it is just the written form of the word." And also: "[F]or Japanese-made words, no matter whether created by semantic transliteration or newly constructed ones, the authorship of creation lies in Japan." Zhou (1998) argues that because Japanese loanwords are written in Chinese characters, this does not reduce their loanword status, but in fact allow them to be borrowed in the first place. Furthermore, orthographically speaking, the argument that any borrowing via Chinese characters does not qualify loanword status was also questioned Xue (2007): "... therefore, similarity in terms of the writing system does not count as an argument for saying Japanese loanwords are not loanwords. The Japanese only adopted Chinese characters to express words from other languages or to create words by themselves based on their own new words."

A relaxed definition of Japanese loanwords can be stated as below:

Relaxed Definition A word is granted loanword status if the following conditions are true:

- either the sense and/or the pronunciation and/or the word form cíxíng (詞形), has been borrowed (at least one aspect of the three),
- the origin of the word can be traced back to a different people and culture
- the loanword is in general use for some time.

Now, because Japanese loanwords have been coined by the Japanese, these words express Japan's reaction to modernization; and because the combination of Chinese characters was previously not part of the Chinese lexicon, or if so, the whole word did not have the modern meaning; ergo, Japanese loanwords are loanwords.

In reaction to the under-appreciation of graphic loans, Tranter (2009) formulates a so-called Graphic Loan Theory. He argues correctly that graphic loans have long been neglected in loanword theories, especially those influenced by the Western tradition. His criticism is based on the argument that in modern literate societies words are exchanged verbally and in written form. Exchanging words across language borders can happen, again, verbally (by phonological reference), by translation and via the borrowing of graphs. We label this process silent borrowing.

An important point of Tranter's theory is to demonstrate that graphic loaning even occurs between languages which do not share the same writing system, e.g. between English and Japanese. He discusses the example of US English 'jitterbug', which was arguably used by GIs stationed in Japan (Tranter 2009, p. 25), and loaned verbally into Japanese, transliterated as /ʤiɾuba/. Normally, we would expect the orthographically visible "tt" also to be transliterated as */ʤit͟taabaggu/ or /*ʤit͟taabagu/ (underlines added for clarity), if—as in most other English loanwords into Japanese—the borrowing would have happened by "eye route" (Tranter 2009, p. 25), that is copying via spelling, not verbally. He uses this counter example to demonstrate that in fact the standard way of borrowing from English into Chinese, Japanese and other Asian languages has happened by spelling, that is by graphic loaning. He further argues that graphic loans deserve more attention, since they are common, systematic and regular.

According to Tranter, graphic loaning happens in a three-step process: introducing a new idea in form of writing (stage 0); the language community produces different phonological realizations of the graph, based on the orthographic and phonological rules of the language community (stage 1); filtered by phonological constraints, the phonological form is then transformed into the script system of the target language (stage 2), compare Table 9 below.

Tranter demonstrates this three-step process in different language settings: twice from English to Japanese (Katagana), from Chinese to Japanese (Katagana), from Chinese to Vietnamese, from Japanese to Korean, from Italian to English, and from Japanese to English. Yet, in all those cases, the target language uses a phonological script system for step 2. Unfortunately, he withholds from us the most interesting example of Chinese as target language. In fact, the idea of graphic loaning expressed in Tranter (2009) is still basically a process of phonological borrowing, but with the difference that the phonological information is not obtained directly via verbal actualization but indirectly via spelling. Spelling introduces deviation errors compared to the actual phonological realization. The theory itself, indeed, incorporates a step of phonological adjustment (step 1); and another step that takes

Table 9 3 step process of Graphic Loans by Tranter (2009)

Stage (Input/Output)	Process
0 SL Forms (US English, dialect) Phonological/Phonetic Input:	[ˈʤirəbʌg] = \<jitterbug\> [ˈʤirəbʌg]
1 Phonological Rules: Phonological Output:	(Can be expressed in terms of Optimality Theory) /ʤiɾuba/
2 Graphic Rules:	Transcription of phonological output in *katakana*: /ʤi/ → < ジ > /ɾu/ → < ル > /ba/ → < バ >
Graphic Output:	< ヅルバ >
TL Forms (Japanese):	/ʤiɾuba/ = < ヅルバ >

this phonological adjustment as input for writing systems (step 2), which in all given examples are again phonological in nature.

A recent compelling approach has been undertaken by Cook (2018). She differentiates between (a) calques, (b) silent borrowing ('symbolic loans', e.g. Japanese loanwords), (c) material borrowing ('transliterations', i.e. borrowing sound-meaning pairs), (d) orthographic loans (cases of borrowing orthography and sound, but not the pronunciation—'graphic loans'), (e) a class which we do not recognize as loanwords, which are labeled 'wholesale loans', that are foreign unaltered words in Chinese; and (f) so-called 'hybrids', which are any words with mixed borrowing properties. Laudable is Cook's differentiation between 'partial calques' (e.g. *daoban* 盗版 'pirate copy') by which each morpheme of the source word is mapped onto the target word, and 'holistic calques' (e.g. *diannao* 電腦 'computer'), by which only the word sense is being borrowed, but differently realized in the target language, a process which we label conceptual remakes (Cook 2018, p.8-9 and footnote 8). Unfortunately, she stops shortly before also recognizing that the so-called partial calques hint towards the more general property of Chinese characters as being able to fully or partially represent sound and meaning of native words as well as of borrowed words, with important ramifications for a theory of loanwords in Chinese. This is a topic we have explored here in terms of isomorphemic and complex transliteration, as well as conceptual remakes.

At this point we want to steer our more historically oriented discussion into the home stretch, aiming at a general classification of loanwords in Chinese. The first purpose of such a classification is to tell loanwords from native words. As Haspelmath (2009) pointed out, for identifying loanwords one needs to be looking for a smoking gun, a similar 'shape and meaning' of words from genealogically different languages. If we take Haspelmath literally and substitute graph for shape, Japanese loanwords immediately fall overboard, simply because their shape and meaning are often very similar, yet they are historically proven loanwords. Hence, some extra-linguistic knowledge will be needed for this task.

The three fundamental principles upon which we build the classification are already present in the very structure of the previous section, namely, the capacity of written Chinese to express sound and meaning separately and selectively due to the use of Chinese characters, that is (a) the principle of representing sound via association of graph (symbol) with word pronunciation; (b) the principle of representing meaning via association of graph with word sense, and (3) the principle of representing a second layer of meaning due to the specific properties of Chinese characters.

Now, let us shortly summarize the state of the art. The following few paragraphs hopefully show the reader where most approaches so far have been insufficient. Firstly, the Chinese literature parts loanwords most often simply into phonetic loans (*yinyi* 音譯) and semantic loans (*yiyi* 意譯). Semantic loans are often equated with calques, either iso-morphemic or complex transliterations, although this distinction is almost never drawn. In fact, most approaches suffice to look only at the word level to draw conclusions about any processes of borrowing. The morphological level is often ignored.

Secondly, a triadic categorization adds hybrid loans, that is loanwords that carry a native suffix, e.g. *-che* (∼車 car) or *-gou* (∼狗 dog). Again, which specific function the suffix and/or the other morphemes fulfill, either towards each other or towards the source word is almost never further investigated.

Scholars often regard Japanese loanwords as a separate class of loanwords altogether, which are labeled, unfortunately, graphic loans in the Western literature, because it was assumed that Japanese Kanji would be part of the elements being borrowed. The necessary adjustments that take place during borrowing point to the fact that the most important element of the copying process is the idea behind the Kanji combination[35], not so much the Kanji themselves, which are in any case written in autochthonous Chinese characters, a process we label silent borrowing instead. The Chinese literature calls those either *hanzici* (漢字詞) as a general term, or *jiexing hanzici* (借形漢字詞, see Shi (2013)), *hezhi hanyu* (和製漢語, see Takeyoshi (1996)), or *huiguici* (回歸詞), if those words are regarded to be of Chinese origin, or *riben jieci* (日本借詞), if those words are regarded as genuine Japanese words not modeled after Western concepts or words.

---

[35] Many Japanese loanwords in Chinese are themselves loanwords from other languages; they are in many cases better understood as mediated loanwords from a third party donor language (Chen 2011; Feng 2016; Pan 2019). For example, Chen (2017) has argued, the term Enlightenment (啟蒙) gained much of its socio-political content in Japan. As such, 'enlightenment' reflects Japan's drive for modernity, openness and Western knowledge. When borrowed into Chinese, it there inherited the traditional connotations of the Chinese expression 'opening up and dispelling ignorance'. It finally merged with the Japanese-mediated term 'movement'. As Enlightenment Movement, it turned into a modern political terminology, a term that showed how to open up the lower echelons of society to the intellectual society of the traditional upper classes, then became a motto for a radical political agenda, quasi an 'Enlightenment by the use of force', and a concept for distinguishing tradition from modernity, and not ignorance from reason.

In terms of orthography, some researchers explicitly mention the addition of a semantic *pianpang*[36] during the process of selecting (or in rare cases, creating) characters during transliteration, such as *jin* (金) for metals in *tai* (鈦 titanium), or *qi* (气) for gases as in *yang* (氧 oxygen) (Liu et al. 1984, Cen 2015, Chu 2006, Fu 2011, Yang 2007). In terms of phonetic loans, Japanese phonetic loans (e.g. *tatami*) are sometimes also explicitly mentioned.

The minimal consensus for any classification of loanwords goes arguably to a set containing just phonetic loans and hybrid loans (complex transliteration in various styles), after Wang (Wang 1958, p. 9) famously rejected the idea of accepting calques as loanwords. He holds on to the Western-influenced principle that loanwords, just as words in general, are phonetic-semantic units, leaving out one aspect or the other from the borrowing process consequently would disqualify the word as loanword. Lü (1992) and Shi (2000, 2019) also follow this principle; however, Shi makes a differentiation whether the calque is a direct transliteration of a Western concept, in which case it counts as a semi-loanword only, or whether it was mediated via Japanese and counts as a graphic loan (this is marked by the asterisk * in Table 10 below). One step further, allowing for a more relaxed definition, scholars would also include some sort of semantic loans (Wang 2015; Masini 1997; Feng 2004; Yang 2007; Zhao 2016). However, in these approaches subtypes of semantic loans are often not clearly separated.

The setup in Table 10 below captures most categories used by most studies of Chinese loanwords. Of course, sometimes we come across extremely elaborated categorizations, such as the ones by Wang (1958), Novotná (1967), mentioned by Masini (1997), Shi (2013), or recently by Cook (2018).

2.4.1 Detailed classification

The following Table 11 lists in detail all most relevant categories of loanwords in Chinese. It is based on a classification by Shi (2013), but modified[37] in two notable ways. We expanded and changed the categories mentioned by Shi, including the category organization, names, examples, explanations, and order. Furthermore,

---

[36] *Pianpang* is a component within a Chinese character that has one of four specific functions: (i) indicating meaning by graph form, as *da* (大 a picture of a person) in *mei* (美), wearing a headdress *yang* (羊 variant graph of a headdress, resembling the form of a sheep antler), together they form here the composite meaning 'person with headdress', which refers to 'beautiful'; (ii) indicating meaning by word sense, as *da* (大 here bringing into plays its modern word sense 'large') in *jian* (尖 pointy) under *xiao* (小 small), together they form here the composite meaning 'small over large', which refers to 'pointy, sharp'; (iii) indicating pronunciation, as *ji* (幾) in *ji* (機), or *shang* (尚) in *tang* (躺); (iv) substituting for more complex forms, as *bei* (貝) in *zei* (賊 thief), whereas *bei* (貝) here substitutes for the phonetic *ze* (貝), now often falsely reinterpreted as semantic component *bei* (則 valueable).

[37] We combined similar categories, e.g. #5 'funny phonetic translations' with #3 'Alliterated Homophones with semantic transliteration' and deleted a few that are not directly concerned with Chinese characters, such as #15 for Chū Nôm characters, and #28, #29 for English, as well as all categories that capture fake loanwords, i.e. Chinese words modeled after loanwords.

Table 10 Simple categorization of loanwords

#	Category	Example	Wang Lida	GR[1]	Liu	Cen	Wang Li	Masini	Feng	Zhu	Yang	Fu	Zhao	Shi	Cook
1	Graphic loans (Jp.)	economy = 經濟	.	✓	✓	.	✓	✓	✓	✓	✓	✓	✓	✓	✓
2	Adding *pianpang*	silicon = 矽	.	.	✓	✓	.	.	.	✓	✓	.	.	.	.
3	Phonetic loans	telephone = 德律風	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
4	Phonetic loans (Jp.)	たたみ = 榻榻米	✓	✓	.	✓	.	✓	.	.	✓	✓	.	✓	✓
5	Calques	hot dog = 熱狗	.	.	.	.	✓	✓	✓	.	✓	.	✓	*	✓
6	Hybrid loans	lotto = 樂透; car = 卡車	✓	✓	✓	.	✓	✓	✓	✓	✓	✓	✓	✓	✓

[1] GR: Gwoyeu Rhybaw Waylaiyeu Tszrdean

Table 11 Detailed classification of Chinese loanwords (based on Shi (2013), modified)

Category	#	Source → Target	Notation
Alphabetic words	1	x_a shares$_b$ → x_a 股$_b$	$[\langle m_a\rangle_S m_b] \to [\langle \mu_a^{(p)c}\rangle_S \langle \mu_b^c\rangle_G]$
		Source word has variable letter or number, semantic transliteration (see p.4).	
	2	$T_a D_b K_c$ cup$_d$ → $T_a D_b K_c$ 杯$_d$	$[\langle s_a\rangle_L \langle s_b\rangle_L \langle s_c\rangle_L \langle m_d\rangle_L] \to$ $[\langle \sigma_a^p\rangle_L \langle \sigma_b^p\rangle_L \langle \sigma_c^p\rangle_L \langle \mu_d^c\rangle_G]$
		Source word has fixed letter, partly abbreviated, Chinese keeps the abbreviation, adds ontological suffix (see p.5).	
	3	$3_a D_b$ effect$_c$ → $3_a D_b$ 效果$_c$	$[\langle m_a\rangle_S \langle s_b\rangle_L \langle m_c\rangle_L] \to$ $[\langle \mu_a^c\rangle_S \langle \sigma_b^{pc}\rangle_L \langle \mu_{c_1}^c\rangle_G \langle \mu_{c_2}^c\rangle_G]$
		Source word starts with a fixed number, partly abbreviated, Chinese keeps the number and abbreviation, partly semantic transliteration (see p.5).	
Numerals, units, scient. notation	4	%$_a$ → %$_a$	$[\langle m_a\rangle_S] \to [\langle \mu_a^c\rangle_S$
		Special character set. Chinese borrows form and meaning, but not its pronunciation (see p.5).	
Graphic loan	5	卍$_a$ → 卍$_a$	$[\langle x_a\rangle] = [\langle \sigma_a\rangle_g^c]$
		Foreign symbol gains new pronunciation in Chinese (see p.6).	
	6	腺$_a$ → 腺$_a$	$[\langle g_a\rangle]_{JP} = [\langle \mu_a^c\rangle_g]_{CN}$
		Chinese borrows Japanese-only Kanji and its word meaning (see p.6).	
	7	鱈$_a$ → 鱈$_{a_1}$魚$_{a_2}$	$[\langle g_a\rangle]_{JP} \to [\langle \mu_{a_1}^c\rangle_g \langle \mu_{a_2}^{c'}\rangle_G]$
		Extended. Source is Japanese-only Kanji, target word adds suffix (see p.6).	
Silent borrowing	8	すし$_{a_{JP}}$ → 寿$_{a_1}$司$_{a_2{JP}}$ → 壽司$_{a_{CN}}$	$[m_a]_{JP} \to [g_{a_1} g_{a_2}]_{JP} \to [\langle \sigma_a\rangle_{g'}\langle \sigma_b\rangle_g]_{CN}^c$
		Chinese borrows Kanji combination, together with the associated word sense, but not its pronunciation (see p.6).	
Silent borrowing, back borrowing	9	(1) 革$_a$命$_{b_{CN}}$ → 革$_a$命$_{b_{JP}}$ (2) 革$_a$命$_{b_{JP}}$ → 革$_a$命$_{b_{CN}}$	(1) $[g_a g_b]_{CN}^c \to [\langle \mu_a^c\rangle_g \langle \mu_b^c\rangle_g]_{JP}^{c'}$ (2) $[g_a g_b]_{JP}^{c'} \to [\langle \mu_a\rangle_g^c \langle \mu_b\rangle_g^c]_{CN}^{c'}$
		(1) Japanese borrows character combination from Chinese, adds meaning shift/extension; (2) Chinese back borrows character combination with new word sense. Characters are adjusted, if necessary (see p.6).	

we added our notation to make explicit the relevant processes that underlie each category.[38]

2.4.2 A Continuum of loanwords in Standard Modern Chinese

The notation system is useful for elucidating various aspects of the borrowing processes and for showing how and why loanword categories differ. But the notation system does not show us, by *how much* they differ. Hence, we might ask

[38] We also harmonized table and text, so that the reader may use the provided page links to find more explanations within the main text.

Table 11 continued

Table 11 – *Continued from previous page*

Category	#	Source → Target	Notation
Silent borrowing, mediated	10	(1) club$_a$ → 倶$_{a_1}$ 楽$_{a_2}$ 部$_{a_3}$ (2) 倶$_{a_1}$ 楽$_{a_2}$ 部$_{a_3}$ → 倶$_{a_1}$ 樂$_{a_2}$ 部$_{a_3}$	(1) $[m_a]_{\text{EN}} \to [\langle \sigma_{a_1}^p \rangle_g^{c'} \langle \sigma_{a_2}^p \rangle_g^{c'} \langle \sigma_{a_3}^p \rangle_g^{c'}]_{\text{JP}}^{c'}$ (2) $[g_{a_1} g_{a_2} g_{a_3}]_{\text{JP}} \to$ $[\langle \sigma_{a_1} \rangle_{g'}^{pc'} \langle \sigma_{a_2} \rangle_{g'}^{pc'} \langle \sigma_{a_3} \rangle_{g'}^{pc'}]_{\text{CN}}^{c'}$
		(1) Phonetic Kanji representation of Western word, the Kanji chosen provide an additional layer of conceptual association; (2) Silent borrowing. Chinese borrows Kanji combination and word sense.(see p.6)	
Phonetic loan	11	sacima$_a$ → 薩$_{a_1}$ 其$_{a_2}$ 馬$_{a_3}$	$[m_a] \to [\langle \sigma_{a_1}^p \rangle_G \langle \sigma_{a_2}^p \rangle_G \langle \sigma_{a_3}^p \rangle_G]$
		Simple phonetic loan. Each syllable in Chinese has phonetic function only (see p.9).	
Phonetic loan, complex orthography	12	から$_a$ オケ$_b$ → 卡$_{a_1}$ 拉$_{a_2}$OK$_b$	$[m_a m_b] \to [\langle \sigma_{a_1}^p \rangle_G \langle \sigma_{a_2}^p \rangle_G \langle \sigma_{b_1}^p \rangle_L \langle \sigma_{b_2}^p \rangle_L]$
		Partly transliterated into Chinese graphs, partly into alphabetic letters (see p.14).	
	13	sir$_a$ → 阿$_{a_1}$ sir$_{a_2}$	$[m_a] \to [\langle \mu_{a_1} \rangle_G \langle \sigma_{a_2}^{pc} \rangle_L]$
		Partly transliterated into Chinese graphs, partly keeps source spelling (see p.14).	
	14	byebye$_a$ → 8$_{a_1}$8$_{a_2}$	$[m_a] \to [\langle \sigma_{a_1}^p \rangle_S \langle \sigma_{a_2}^p \rangle_S]$
		Non-Hanzi orthography and phonetic imitation (see p.14).	
	15	T$_a$-Shirt$_b$ → T$_a$-恤$_b$	$[\langle s_a \rangle_L \langle m_b \rangle_L] \to [\langle \sigma_a^p \rangle_L \langle \sigma_b^p \rangle_G]$
		Phonetic transliteration, partly imitating source orthography (see p.14).	
Phonetic, exchanged	16	Zebra$_a$ → 芝$_{a_1}$ 不$_{a_2}$ 拉$_{a_3}$ → 斑$_{a_1}$ 馬$_{a_2}$	$[m_a] \to [\langle \sigma_{a_1}^p \rangle_G \langle \sigma_{a_2}^p \rangle_G \langle \sigma_{a_3}^p \rangle_G] \to$ $[\langle \mu_{a_1}^{c'} \rangle_G \langle \mu_{a_2}^c \rangle_G]$
		Phonetic replaced by semantic transliteration (see p.10).	
Phonetic, adjusted	17	Italy$_a$ → 義$_{a_1}$ 大$_{a_2}$ 利$_{a_3}$ → 意$_{a_1}$ 大$_{a_2}$ 利$_{a_3}$	$[m_a] \to [\langle \sigma_{a_1}^p \rangle_G \langle \sigma_{a_2}^p \rangle_G \langle \sigma_{a_3}^p \rangle_G] \to$ $[\langle \sigma_{a_1}^p \rangle_{g'} \langle \sigma_{a_2}^p \rangle_G \langle \sigma_{a_3}^p \rangle_G]$
		Phonetic transliteration orthographically adjusted (see p.11).	
Phonetic, mediated	18	Engine$_a$ → エン$_{a_1}$ ジン$_{a_2}$ → 引$_{a_1}$ 擎$_{a_2}$	$[m_a] \to [\langle \sigma_{a_1}^p \rangle_{\text{KG}} \langle \sigma_{a_2}^p \rangle_{\text{KG}}]_{\text{JP}} \to$ $[\langle \sigma_{a_1}^p \rangle_G \langle \sigma_{a_2}^p \rangle_G]_{\text{CN}}$
		Mediated phonetically via non-Hanzi script into Chinese (see p.11).	
	19	天ぷら$_a$ → thian$^{35}_{a_1}$ pu$^{55}_{a_2}$ lah$^3_{a_3}$ → 甜$_{a_1}$ 不$_{a_2}$ 辣$_{a_3}$	$[m_a] \to [\sigma_{a_1}^p \sigma_{a_2}^p \sigma_{a_3}^p]_{\text{MN}} \to$ $[\langle \sigma_{a_1}^p \rangle_G \langle \sigma_{a_2}^p \rangle_G \langle \sigma_{a_3}^p \rangle_G]_{\text{CN}}$
		Mediated from source orally via a dialect (see p.12).	
Iso-morphemic transliteration	20	black$_a$board$_b$ → 黑$_a$ 板$_b$	$[m_a m_b] \to [\langle \mu_a^c \rangle_G \langle \mu_b^c \rangle_G]$
		Semantic mapping of each morpheme (see p.13).	
Complex transliteration, functionally separated	21	car$_a$ → 卡$_{a_1}$ 車$_{a_2}$	$[m_a] \to [\langle \sigma_{a_1}^p \rangle_G \langle \mu_{a_2}^c \rangle_G]$
		Phonetic syllable plus autochthonous suffix (see p.14).	
	22	motor$_a$cycle$_b$ → 摩$_{a_1}$ 托$_{a_2}$ 車$_b$	$[m_a m_b] \to [\langle \sigma_{a_1}^p \rangle_G \langle \sigma_{a_2}^p \rangle_G \langle \mu_b^{c'} \rangle_G]$
		Functionally separated morphemes. At least one source morpheme is semantically substituted (*lun* 輪 'cycle'$_b$ → *che* 車 'car$_b$', see p.14).	
Complex transliteration, expanded	23	ink$_a$ → 墨$_{a_1}$ 水$_{a_2}$	$[m_a] \to [\langle \mu_{a_1}^c \rangle_G \langle \mu_{a_2}^{c'} \rangle_G]$
		Semantic morpheme plus autochthonous suffix (see p.15).	
	24	sir$_a$ → 阿$_{a_1}$ 蛇$_{a_2}$	$[m_a] \to [\langle \mu_{a_1}^{c'} \rangle_G \langle \sigma_{a_2}^p \rangle_G]$
		Prefix combined with a phonetic syllable (see p.14).	

Table 11 continued

<div align="center">Table 11 – <i>Continued from previous page</i></div>

Category	#	Source → Target	Notation
Complex transliteration, fully fused	25	čapan$_a$ → 袷$_{a_1}$ 袢$_{a_2}$	$[m_a] \rightarrow [\langle \mu_{a_1}^{pc'} \rangle_{\mathrm{G}} \langle \mu_{a_2}^{pc'} \rangle_{\mathrm{G}}]$
		Fully fused phonetic and semantic function for each morpheme (see p.14).	
	26	w$_a$w$_b$w$_c$ → 萬$_a$ 維$_b$ 網$_c$	$[\langle s_a \rangle_{\mathrm{L}}^{c} \langle s_b \rangle_{\mathrm{L}}^{c} \langle s_c \rangle_{\mathrm{L}}^{c}] \rightarrow$ $[\langle \mu_a^p \rangle_{\mathrm{G}}^{c'} \langle \mu_b^p \rangle_{\mathrm{G}}^{c'} \langle \mu_c^p \rangle_{\mathrm{G}}^{c}]^{c'}$
		Fully fused phonetic and semantic function for each morpheme, based on abbreviated source word. Semantic functions of the graphs add a layer of conceptual reinterpretation of the word sense (see p.16).	
Complex transliteration, partially fused	27	shoot$_a$ → 述$_{a_1}$ 球$_{a_2}$	$[m_a] \rightarrow [\langle \sigma_{a_1}^p \rangle_{\mathrm{G}} \langle \sigma_{a_2}^p \rangle_{\mathrm{G}}^{c'}]^{c'}$
		Phonetic transliteration, but semantic function via graph (see p.14).	
	28	khatan$_a$ → 卡$_{a_1}$ 墊$_{a_2}$	$[m_a] \rightarrow [\langle \sigma_{a_1}^p \rangle_{\mathrm{G}} \langle \mu_{a_2}^{pc} \rangle_{\mathrm{G}}]$
		Partially fused phonetic and semantic function for some units (see p.16).	
	29	mini$_a$skirt$_b$ → 迷$_{a_1}$ 你$_{a_2}$ 裙$_b$	$[m_a m_b] \rightarrow [\langle \sigma_{a_1}^p \rangle_G^{c'} \langle \sigma_{a_2}^p \rangle_G^{c'} \langle \mu_b^c \rangle_G]^{c'}$
		Units functionally separated. Morphemes with phonetic function have graphs which add a second layer of meaning (see p.16).	
Conceptual remakes	30	objective$_a$ → 客$_{a_1}$ 觀$_{a_2}$	$[m_a] \rightarrow [\langle \mu_{a_1} \rangle_{\mathrm{G}} \langle \mu_{a_2} \rangle_{\mathrm{G}}]^{c'}$
		New semantic expression of the source word sense (see p.16).	
Sense extension	31	human$_a$rights$_b$ → 人$_a$ 權$_b$	$[m_a m_b] \rightarrow [\langle \mu_a^c \rangle_{\mathrm{G}} \langle \mu_b^c \rangle_{\mathrm{G}}]$
		Semantic transliteration, using semantically extended suffixes (see p.17).	

how to transform the qualitative differences into quantifiable differences to conceptualize a numerical continuum of loanword categories. In fact, the question of how much loanword categories differ from each other can be expressed in two ways: (a) to which degree is a loanword different from a native Chinese word, in terms of orthographic, phonetic and semantic integration, and (b) to which degree is a loanword different from its respective source word, again in terms of orthography, phonology and semantics. The first question is about the degree of linguistic integration and a measurement of how much effort a language community has put into making the loanword look and feel like a standard native word. The second question is about some sort of distance between source and target word, i.e. how much the target words has been reshaped in comparison to its original source word.[39]

Both dimensions create the conceptual space within which we measure and compare every loanword category, and in fact each loanword. Each dimension ('Degree of Integration' and 'Distance') is calculated by assigning specific values to four parameters[40]: (a) unit: syllable or morpheme; (b) function of each unit:

---

[39] Although we try a numerical approach, the calculation is supposed to be understood as a vehicle for comparing categories, in lack of a better approach.

[40] The values are in and by itself *not* reflective of certain linguistic properties, but are designed to facilitate a meaningful numerical comparison.

phonetic, semantic, or both; (c) word length; (d) orthography: Chinese characters or alphabetic letters, also considering if phonetic and semantic functions are outsourced from morpheme to graph. For 'Integration', the value for each loanword is measured against a typical Chinese word; for 'Distance', it is compared to the values of the respective source word.

To give a slightly simplified example, source word 'car; (truck)', is rewritten as $<mK - pc > L$[41], target word *kache* '卡车', written as $<s.P[i;b] > G, <m.K[b;t] > G$[42], whereas after assigning specific conceptual values to each parameter, we compute two values between 0 and 1 (for the target word only): Degree of Integration = 0.71, Degree of Distance = 0.52.

The higher the degree of integration, the more similar the word to a native Chinese word. The higher the value for distance, the greater the difference between source and target words. For each loanword category, we computed a various amount of words.[43] We used k-means for clustering.

If plotted with 'Distance' as the x-axis and degree of 'Integration' on the y-axis, loanword categories form three clusters: (a) Group 1 in the top left sector: high integration—in many cases not easily recognizable as loanwords. (b) Group 2 in the right sector—'half-way' integration (with the exception of category 7), generally far detached from their original sources—the words have been transformed, but not very successfully so. They are all still to some degree 'loanwordy', either due to a use of non-Hanzi orthography or by a non standard use of morphemic functions. (c) Group 3 in the bottom sector: these words are not very well integrated (which does not contradict high usage), they still feel and look very 'loanwordy', either because they contain alphabetic elements or exhibit a non-standard morphosyntactic functionality. Some of these words are probably going to be substituted by better integrated words over time (Fig. 1).

We think some cautious generalizations about the relationship between these clusters of loanwords can be drawn: (a) loanwords can change their category; e.g. going from phonetic towards a more semantic approach. This has been widely observed. This change would be reflected in an upward move on the continuum; (b) the more effort Chinese makes in transforming a loanword—by selecting native syllables, morphemes and graphs—the more the loanword shifts to the right side of the continuum, increasing the distance from its source word. This right-handed movement does not necessarily lead to integration though; (c) the best integrated loanwords are words that follow Chinese morphosyntactic rules, especially fusing phonetic and semantic functions, either by selecting specific morphemes or by a

---

[41] Here, m stands for morpheme, K stands for combining phonetic and semantic function, p stands for sound and c for meaning, the breakets < > indicate morpheme boundaries, L stands for orthographic representation in the Latin character set.

[42] The source word 'car' is expressed in the loanword as one syllable s and one morpheme m; the division of labour is as follows: the syllable s has no meaning inside the word, hence works only as a phonetic P, related to the source word by representing its sound i but not its meaning b, written in G; m actuates its phonetic and semantic function K, but is related to its source in terms of semantics t, but not phonetics b, expressed in a graph G. Our actual formulas capture slightly more detail.

[43] Since not all categories have an equal amount of members, we think that the computed values are better understood as a rough conceptual estimate and not as an accurate quantitative assessment.
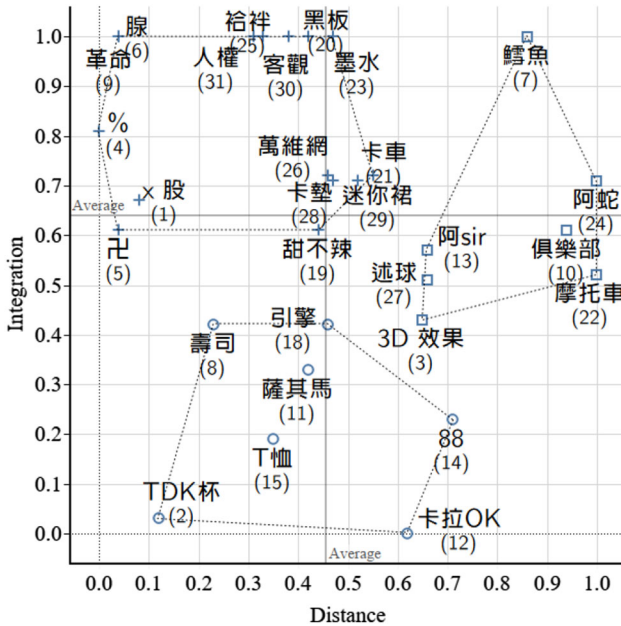
Fig. 1 Continuum of loanword categories

division of labor between syllable and graph, making use of the vast quantity of homonyms and the capacity of Chinese characters to selectively express sound and meaning.

3 Summary

This paper gives a general overview about loanwords in Modern Standard Chinese. Based on the fundamental functions of written language, we differentiate between borrowing orthographic, phonetic and semantic properties. We argue that alphabetic words, numerals, units of measurement and foreign graphs have been neglected under the ordinary definition of loanwords; hence, we have included them under orthographic borrowing. Due to large numbers of homonyms, characters can express phonetic and semantic values selectively; therefore processes of borrowing into Chinese are not trivial. Borrowing from Japanese into Chinese often triggers an exchange of set-equivalent variant characters. Phonetic borrowings are very common and most recognized, but they also can be updated over time by alternating one or more graphs of a word. In some cases the entire borrowing strategy can switch from phonetic to semantic or complex transliteration.

In terms of semantic borrowing, questions may arise that are unknown in Western languages. A written Chinese word has not only a semantic value (referent) associated with its entire word meaning, but also an additional second layer of

meaning that is created due to the literal reading of the semantic function of each character. Each character always has at least one semantic value associated with itself (the pictographic meaning or 'root meaning' of the character) which may or may not be independent of the modern word sense in which the character takes part. This fact largely increases the—often playfully exploited—ability to create ambiguity.

We define loanwords by virtue of the three fundamental mapping principles that are relevant during the borrowing process between any source language and Chinese: (a) the principle of representing sound via association of graph (symbol) with word pronunciation; (b) the principle of representing meaning via association of graph with word sense, and (c) the principle of representing a second layer of meaning due to the use of Chinese characters. This investigation also includes, on a morpheme level, observing how the subparts of a Chinese word functionally relate to each other, to the entire word meaning, and to each counterpart of the source word. Based on this thorough assessment, we summarize our findings in a detailed classification of loanwords.

With great detail often comes confusion; however, the devil is in the details. Why, for example, do we regard *julebu* (俱樂部) as silent borrowing, not a phonetic loan in relation to 'club', or even as some flavor of a semantic borrowing process, since all three characters read together may trigger a different meaning? Here extra-linguistic knowledge is needed to decide whether a factual borrowing instance may or may not conflict with linguistic principles. An even more difficult case is the differentiation between phonetic loans (with complex sense, Category 19) and complex transliteration (Categories 27–33). For example, *leida* (雷$_a$達$_b$ radar) is considered to be a complex transliteration after a careful weighting of the triadic relationship mentioned earlier. However, it could be argued that neither thunder$_a$ nor reach$_b$ are in any shape or form semantically associated with the word sense of radar, hence *leida* could be safely considered phonetic. Yet, we reply to this argument that, given the phonological mapping constraints, thunder$_a$ and reach$_b$ are not supposed to be directly associated with the word sense 'system of detecting the presence of objects by radio waves', yet somehow 'thunder' encapsulates a metaphorical sense of 'wave' and 'reach' probably some sort of 'detecting'. There surely is room for discussion. A definite answer can only be given, if we are able to trace back the very first attested occurrences in all relevant languages and make a decision based on time and date.

In terms of the theoretical development, most Chinese scholars today acknowledge the importance of Japanese loanwords, which has softened their standpoint on calques and semantic loanwords in general, but there exist exceptions nonetheless. The times that a loanword had to be phonetic—as earlier Western-influenced Chinese scholars argued—are over. In recent years, interesting theoretical developments have been undertaken to harmonize Western loanword theories with Chinese theories. Tranter's Graphic Loan Theory, Shi's investigation into Japanese loanwords and Cook's research, among others, have definitely benefited the field. Future research should take on the problem of compound words in Chinese. Currently, compound words are excluded by principle from any loanword status, influenced by Western theories, but in Chinese, compounding is a strong force in the

development of new words, as well as in integrating loanwords and must not be excluded. Rather, compounding can trigger interesting phenomena on the morpheme level, especially when they are created by intent, which happens during active loanword integration. Hence, our senses should be alerted when looking out for selective use of orthographic, phonetic and semantic functions that occur during any mapping between source and target words.

Although Chinese has doubtless the longest written history of the world, traceable impact of neighboring languages—with the exception of Japanese—upon its lexicon is astonishingly limited. Now, Japan on the other hand, due the course of history, was at least 30 years ahead in terms of absorbing and processing Western knowledge, before the Chinese started to massively borrow from the Japanese lexicon. The argument has often been brought forward that Japanese loanwords are not loanwords because they are written in Chinese characters and 'feel Chinese'; hence, they could just as well have been made in China. By introducing the concept of silent borrowing, opposed to early 'graphic loaning', we tried to elucidate all relevant processes during this specific type of borrowing and can conclude that loanwords from Japan count as loanwords. Even more so, we forward the argument that a considerable number of Japanese loanwords are of Western origin, created by Westerners in China before the Meiji Restoration began.

Finally, after a detailed categorization of all major loanword categories and many minute details between them, we sum up this investigation by providing a large picture. We map the loanword categories provided in Table 11 on a conceptual 2-dimensional space, measuring for any category (a) by how much a loanword differs from a standard native Chinese word ('Degree of Integration'), and (b) by how much a loanword is different to its respective source word ('Distance'). This measurement is done for various numbers of loanwords in each category. Since different classes have a different count of members, this calculation is better understood as a conceptual assessment rather than an exact numeric evaluation. However, by doing so, we hopefully exemplify not only how loanword categories relate to each other, but also the kind of relationship that initiates a loanword into the Chinese lexicon in general.

References

Baxter, W.H., and L. Sagart. 2014. *Old Chinese: A new reconstruction*. Oxford: Oxford University Press.
Booij, G., and J. Van Marle. 2001. *Yearbook of morphology 1999*, vol. 9. New York: Springer.
Cen, Q. 2015. *[Chinese loanword dictionary] Hanyu wailaici cidian*. Beijing: Commercial Press.

Chen, L. 2011. "[The Formation of Modern Chinese and Japanese Concepts and Their Interaction] Jindai zhong ri gainian de xingcheng ji qi xianghu yingxiang". In: Dongya guannian shi jikan bianshen weiyuanhui ta bian "dongya guannian shi jikan" 1, pp. 149–78.

Chen, K.-J. et al. 1996. "Sinica corpus: Design methodology for balanced corpora". In: Proceedings of the 11th Pacific Asia Conference on Language, Information and Computation, pp. 167-176.

Chen, J.-s. 2017. "[When Enlightment Meets "Qimeng Yundong": The Translation of "Enlightment" in Bilingual Dictionaries] Shuangyu cidian yu ciyuan kaosuo: yi qimengyundong wei li de taolun". In: Intellectual History. Taipei: Lianjing Chubanshe.

Chu, C.-N. 2006. [Research on Chinese Linguistics for Fifty Years] Wushi nianlai de zhongguo yuyan xue yanjiù. Taiwan: Xuesheng Shuju.

Cook, A. 2018. A typology of lexical borrowing in Modern Standard Chinese. In Lingua Sinica 4.1, 6.

Ding, H. et al. 2017. A Preliminary Phonetic Investigation of Alphabetic Words in Mandarin Chinese. In INTERSPEECH, 3028–3032.

Feng, T. 2004. [Origin of New Languages: Cultural Interaction between China, the West, and Japan and the Generation of Modern Chinese Character Terms] Xinyutanyuan: zhongxiri wenhua hudong yu jindai hanzi shuyu shengcheng. Zhonghua Chubanshe.

Feng, T. 2016. [A study of the generation and evolution of modern terms in Chinese characters and the cultural interaction among China, the West and Japan] Jindai hanzi shuyu de shengcheng yanbian yu zhong xi ri wenhua hudong yanjiu. Beijing: Jingji Kexue Chubanshe.

Foley, T. S. 2009. Biblical translation in Chinese and Greek: Verbal aspect in theory and practice. Vol. 1. Brill.

Fu, Z. 2011. [Modern Chinese Lexicon] Xiandai hanyu cihui. Hong Kong: Commercial Press.

Haspelmath, M. 2009. Lexical borrowing: Concepts and issues. Loanwords in the world's languages: A comparative handbook 35–54.

Haspelmath, M., and U. Tadmor. 2009. Loanwords in the world's languages: a comparative handbook. Berlin: Walter de Gruyter.

Haugen, E. 1972. Semicommunication: The language gap in Scandinavia. I: The ecology of language. Essays by Einar Haugen. Selected and introduced by Anwar S. Dil. In: Stanford, Ca.[Tidigast tryckt 1966] 215.

Huang, C.-R. 2009. Semantics as an orthography-relevant level for Mandarin Chinese. In The 17th Annual Conference of the International Association of Chinese Linguistics, Paris, July 2–4 2009.

Huang, C. and H. Liu. 2017. Corpus-based automatic extraction and analysis of Mandarin alphabetic words. In Journal of Yunnan Teachers University. Philosophy and social science section.

Kang, Y. 2011. Loanword phonology. In The Blackwell companion to phonology, 1–25.

Li, Y. and B. Qian. 1998. [Complete Works of History of Science] Li Yan Qian Baozong kexue shi quanji. Ed. by S. Du, S. Guo, and L. Dun. Liaoning Jiaoyu Chubanshe.

Li, X., and Z. Yu. 2005. [A Study of Variant Characters in Modern Chinese] Xiandai hanyu yuxingci yanjiu. Shanghai: Cishu Chubanshe.

Liu, S. (1900 (2018)). [Illustration of the Word Class in Cheng Ai Meng School] Cheng ai meng xuetang zike tushuo. Taipei: Yuanshen Chubanshe.

Liu, L. 2008. [Interlingual Practice: Literature, National Culture and Translated Modernity - China, 1900-1937], kuayuji shijian: wenxue, minzu wenhua yu bei yi jie de xiandaixing. Trans. by S. W. et al. Beijing: Shenghuo Dushu Xinzhi San Lian Shu Dian.

Liu, Z., et al. 1984. [Chinese Loanword Dictionary] Hanyu wailaici cidian. Shanghai: Shanghai Cishu Chubanshe.

Lü, S. 1992. Outline of Chinese Grammar, vol. 1. Beijing: Commercial Press.

Masini, F. 1997. [The formation of modern Chinese lexicon and its evolution toward a national language: the period from 1840 to 1898] Xiandai hanyu cihui de xingcheng: shijiu shiji hanyu wailaici yanjiu. Trans. by H. Huang.

Novotná, Z. 1967. Contributions to the study of loan-words and hybrid words in modern Chinese. Archiv Orientalni 35: 613–648.

Pan, G. 2019. [Creating World Knowledge in Modern China] Chuangzao jindai zhongguo de shijie zhishi. Beijing: Shehui Kexue Wenxian Chubanshe.

Paradis, C. and A. Tremblay. 2009. Nondistinctive features in loanword adaptation. In C. a. Wetzels (Ed.), Loan Phonology. Current Issues in linguistics Theory, 307, 211–224.

Peperkamp, S., I. Vendelin, and K. Nakamura. 2008. On the perceptual origin of loanword adaptations: experimental evidence from Japanese. Phonology 25 (1): 129–164.

Pu, Z. 2001. [The sheep culture and lucky symbols of the Yi people] Yizu yang wenhua yu jifu. Yunnan Renmin Chubanshe.

Qu, L. 1973. *[Studying abroad in the Late Qing] Qingmo liuxue jiaoyu*. Taipei: Sanmin Shuju.

Schmidt, C. and J.-s. Chen (Unpublished). "Silent Revolution—The Western contribution to the Chinese lexicon". Forthcoming.

Shen, G. 2017. [A Cycle of Modern English-Chinese Dictionaries: From W. Lobscheid through Inoue Tetsujiro to the Commercial Press]. In *Intellectual History*. Taipei: Lianjing Chubanshe, 63–102.

Shi, Y. (ed.). 2019. *[Xinhua Loanword Dictionary] Xinhua wailaici cidian*. Beijing: Commercial Press.

Shi, Y. 2000. *[Loanwords in Chinese] Hanyu wailaici*. Changchun: Jilin Jiaoyu Chubanshe.

Shi, Y. 2013. *[Chinese loanwords] Hanyu wailaici*. Beijing: Commercial Press.

Shi, X., Y. Lin, and X. Liyan. 1989. *[Eternal mystery] Qiangu zhi mi*. Zhongzhou: Guji Chubanshe.

Takeyoshi, S. 1996. ["Chinese words of Japanese origin" Chinese Encyclopedia Dictionary] "Hezhi hanyu" hanyu baike dacidian. Meiji Shoin.

Tan, H. 2018. *[History of Japanese Study Abroad in China in Modern Times] Jindai riben dui hua guan pai liuxue shi*. Beijing: Shehui Kexue Wenxian Chubanshe.

Tian, S. 2005. *[Westernization of Chinese Mathematics] Zhongguo shuxue de xihua licheng*. Jinan: Shandong Jiaoyu Chubanshe.

Tranter, N. 2009. Graphic loans: East Asia and beyond. *Word* 60 (1): 1–37.

Wang, L. 2015. [History of the Chinese lexicon] Hanyu cihui shi. Zhonghua Shuju.

Wang, L. 1958. [Borrowed Words from Japanese in Modern Chinese] Xiandai hanyu zhong cong riyu jie lai de cihui. *Zhongguo Yuwen* 2: 90–94.

Wang, L. 1980. *[History of Chinese] Hanyu shigao*. Beijing: Zhonghua Shuju.

Weinreich, U. 1953. Languages in contact. In *The Hague: Mouton*.

Wiebusch, T. and U. Tadmor. 2009. Loanwords in Mandarin Chinese. In 575–598.

Xu, Z. 2013. *[Dictionary of Lianmianci] Lianmianci da cidian*. Beijing: Commercial Press.

Xu, P., and H. Zhang. 2006. *[History of Mathematical Symbols] Shuxue fuhao shi*. Beijing: Kexue Chubanshe.

Xue, C. 2007. *[Language contact and language comparison] Yuyan jiechu yu yuyan bijiao*. Shanghai: Shiji Chubanshe.

Yang, X. 2007. *[Research on Chinese Loanwords] Hanyu wailaici yanjiu*. Shanghai: Renmin Chubanshe.

Zeng, X. 2017. *[A Study of Loanwords in Modern Chinese under the Theory of Complex Dynamic Systems] Fuza dongtai xitong lilun xia de xiandai hanyu wailaici yanjiu*. Guangzhou: Shijie Tushu Chubanshe.

Zhang, L., L. Ni, and L. Pan, eds. 2018. [Cantonese Macao Dictionary] Xianggang aomen da cidian. Tiandi Tushu Chubanshe.

Zhang, Y. 1958. [Can there be so many Japanese loanwords in Mandarin Chinese?] Xiandai hanyu zhong nengyou zheme duo jieci ma? *Zhongguo yuwen* 6: 299–299.

Zhao, M. 2016. *[A Study of Chinese Loanwords in Ming and Qing Dynasties] Ming qing hanyu wailaici shi yanjiu*. Xiamen: Xiamen University Press.

Zhou, J. 1995. [Outline of Chinese vocabulary research history] Hanyu cihui yanjiu shigang.

Zhou, Y. 1998. [On loanwords from Japanese loanwords from Chinese origins via Chinese characters - Japanese loangraphs] Lun you hanzi dailai de hanyu riyuan wailaici - riyu jiexingci. *Hanzi wenhua* 4: 18–22.